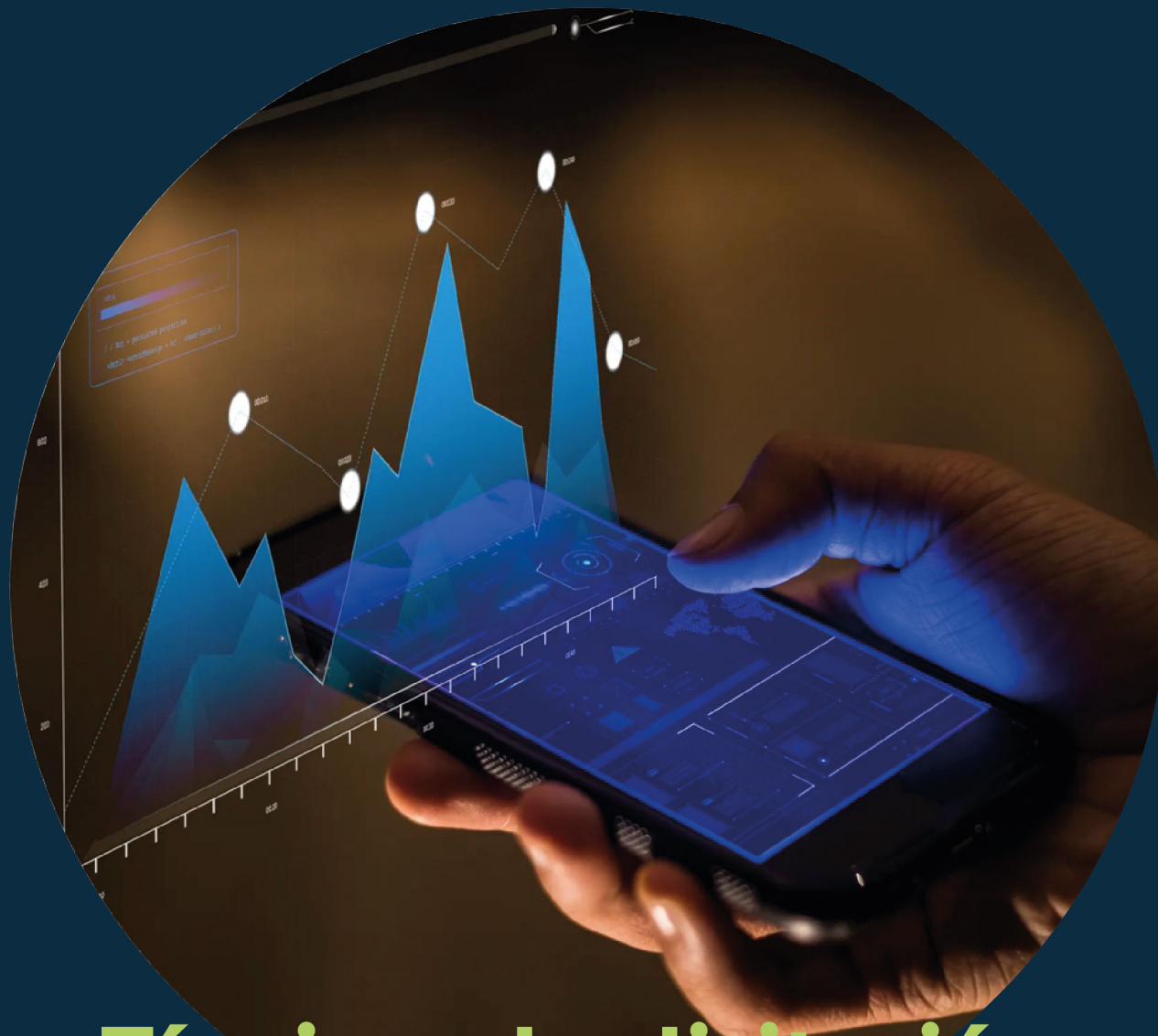




Institución
Universitaria | 80
Reacreditada en Alta Calidad | Años



Editorial
ITM | 25
Años



Técnicas de elicitación de información personal para su uso en estadística bayesiana

Juan Carlos Correa Morales • Carlos Javier Barrera Causil

TÉCNICAS DE ELICITACIÓN DE
INFORMACIÓN PERSONAL PARA SU USO EN
ESTADÍSTICA BAYESIANA

TÉCNICAS DE ELICITACIÓN DE INFORMACIÓN PERSONAL PARA SU USO EN ESTADÍSTICA BAYESIANA

Juan Carlos Correa Morales

Escuela de Estadística, Facultad de Ciencias, Universidad Nacional de Colombia,
Sede Medellín

Carlos Javier Barrera Causil

Facultad de Ciencias Exactas y Aplicadas, Institución Universitaria ITM



Colección Línea Profesor
Correa Morales, Juan Carlos, autor. | Barrera Causil, Carlos Javier, autor. |
Técnicas de elicitación de información personal para su uso en estadística bayesiana / Juan Carlos Correa Morales,
Carlos Javier Barrera Causil (autores). Medellín : Institución Universitaria ITM, Editorial ITM 2025. | Primera
edición. XIV, 218 páginas ; 21.5 x 27.9 cm. | Ilustraciones.

1. Estadística bayesiana | 2. Elicitación de información | 3. Métodos estadísticos | I. Tít. II. Serie

519.542: Estadística bayesiana

Primera edición: abril de 2025

Juan Carlos Correa Morales, Carlos Javier Barrera Causil (autores).

© Institución Universitaria ITM
Sello Editorial ITM
Calle 75 75-101
Medellín, Colombia
Teléfono: 604 440 51 00 ext. 5197
<http://catalogo.itm.edu.co>
editorialitm@itm.edu.co

ISBN DIGITAL: 978-628-7751-18-7

DOI: <http://doi.org/10.22430/reporte.6684>

Corrección de estilo

Martha Cecilia Caballero Jerez

Diseño y diagramación en LaTeX

Juan Carlos Correa Morales
Carlos Javier Barrera Causil

Diseño de cubierta

Mauricio Raigosa Álvarez

Institución Universitaria ITM | Vigilada Mineducación. Reconocimiento de carácter académico: Resolución 6190 del 21 de diciembre de 2005, Mineducación. Reconocimiento de personería jurídica: Decreto 180 del 25 de febrero de 1992, Minjusticia. Renovación acreditación institucional de alta calidad, 8 años: Resolución 013595 del 24 de julio de 2020, Mineducación.

Las ideas y opiniones de este libro son responsabilidad exclusiva de los autores, quienes son igualmente responsables de las citas, referencias y de la originalidad de su obra. En consecuencia, el ITM no responderá ante terceros por el contenido técnico o ideológico del texto ni asume responsabilidad alguna por las infracciones a las normas de propiedad intelectual. Todos los derechos reservados. El texto puede ser reproducido en todo o en parte y por cualquier medio citando la fuente.

What parts are played by experience, by theory, and by imagination in the creation of a system? There is no definite answer. The importance of experience and of theoretical knowledge is not always clearly perceived. Imagination seems a gift, a godsend or the result of beneficial heredity. But is not imagination in fact the result of maturation of knowledge gained during educational and professional practice? Is it not born from facts apparently remembered when circumstances call them back? Is not imagination partly based on the ability to connect notions which, at first sight, look quite unrelated, such as mechanics, electronics, optics, foundry, and data processing—to catch barely seen analogies—like Alice in Wonderland, to go “through the looking glass”?

Gerald Farin

Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide

Índice general

Índice de figuras	V
Índice de tablas	VII
Prefacio	VIII
Introducción	X
1. Probabilidad subjetiva o personal	1
1.1. Probabilidad	1
1.2. Clasificación de las distribuciones a priori	3
1.3. Distribuciones a priori no informativas	4
1.3.1. Distribuciones a priori informativas	4
1.3.2. Probabilidad personal	5
1.4. Probabilidad subjetiva, apuestas y loterías	5
1.5. Probabilidad personal o subjetiva	8
1.5.1. Grados de creencia	10
1.6. El sujeto racional	10
1.6.1. Supuestos de coherencia	11
2. Elicitación	14
2.1. Fuentes de información para la construcción de a priori informativas .	16
2.1.1. Estudios previos similares	17
2.1.2. Opinión de expertos	17
2.2. Elicitación a partir de expertos	19
2.2.1. Tipos de expertos	19
2.2.2. El proceso de elicitación	20
2.2.3. Entrevista personal	21
2.2.4. Problemas con los expertos	22
2.2.5. Sobre los expertos	29
2.2.6. Proceso de elicitación	29
2.2.7. Fallas en el proceso de elicitación	33

3. Métodos de elicitación usando expertos	37
3.1. Formas de elicitación	38
3.2. Métodos gráficos	39
3.2.1. Gráficos de distribuciones de probabilidad	39
3.3. Algoritmo para la determinación de la a priori paramétrica	43
3.4. Métodos directos	51
3.5. Obtención directa de probabilidades	52
3.5.1. Histogramas	52
3.5.2. Método de intervalo fijo	52
3.5.3. Método de intervalo variable con particiones equiprobables o método de bisección	52
3.5.4. Función de distribución acumulada	56
3.5.5. Función densidad de probabilidad	57
3.6. Comparación de métodos de elicitación	59
3.7. Técnicas de elaboración de preguntas y entrevistas para la elicitación	59
3.8. Métodos basados en apuestas	60
3.8.1. Limitaciones del método de apuestas para elicitar	61
3.9. Problemas con la elicitación	61
3.9.1. Peligros en el proceso de elicitación	62
3.9.2. Parámetros poblacionales a elicitar	62
3.10. Validación del proceso de elicitación	63
3.11. Automatización del proceso de elicitación	64
3.12. Conclusiones y recomendaciones	65
4. Método delphi	66
4.1. Pasos de un delphi	67
4.2. Un modelo para el proceso de elicitación	68
4.2.1. Preliminares	68
4.2.2. Elicitación	69
4.2.3. Análisis de los juicios subjetivos	70
4.2.4. Sobre los cuestionarios	71
4.2.5. Ejecución de la segunda vuelta	72
4.2.6. Errores en un delphi	74
4.3. Combinación de la información de múltiples expertos	75
4.3.1. Análisis exploratorio de expertos	81
4.3.2. Conclusiones y recomendaciones	83
5. Aplicaciones	85
5.1. Elicitación de la media	85
5.1.1. Familia conjugada normal: media y precisión desconocidas	90
5.2. Métodos indirectos	97
5.2.1. Apuestas	98
5.2.2. Método de verosimilitudes relativas	98

5.2.3.	Previsión de muestras futuras	98
5.3.	Apuestas	117
5.3.1.	Un diálogo para la obtención de la proporción	118
5.3.2.	Problemas con el uso de loterías	123
5.4.	Elicitación en el modelo lineal	124
5.4.1.	El algoritmo	124
5.4.2.	Ilustración	127
5.5.	Ejemplo hipotético	129
5.6.	Elicitación en el modelo lineal generalizado	138
5.6.1.	Componentes del <i>MLG</i>	138
5.6.2.	Elicitación en el modelo logístico	139
5.6.3.	Uso de muestras hipotéticas	139
5.6.4.	Determinación de la edad de la menarquia	140
5.7.	Elicitación de la distribución a priori informativa del coeficiente de correlación	144
5.7.1.	Conclusiones y recomendaciones	145
6.	Elicitación a partir de información disponible de un solo estudio	146
6.1.	Uso de estudios previos para la determinación de la distribución a priori	146
6.2.	Elicitación de los parámetros de una a priori cuasiconjugada para una población	147
6.3.	Elicitación de una proporción	148
6.4.	Elicitación de los parámetros de una distribución multinomial	149
6.4.1.	Problema de un torneo de tenis	151
6.5.	Elicitación del parámetro de un modelo de Poisson	153
6.6.	Elicitación de los parámetros del modelo lineal	154
6.7.	Elicitación de los parámetros de una normal multivariable	155
6.7.1.	Distribución normal multivariable	155
6.7.2.	Distribuciones condicionales	156
6.7.3.	Distribución de Wishart	156
6.7.4.	Distribución <i>t</i> multivariable	156
6.7.5.	Familia conjugada de los parámetros de la población normal multivariable $N(\mathbf{m}, \mathbf{R})$	157
6.7.6.	Elicitación de los parámetros de la a priori conjugada	165
6.8.	Distribución a priori potencia	167
6.9.	Elicitación de los parámetros del modelo lineal	168
6.9.1.	Ilustración	168
6.10.	Elicitación de los parámetros del modelo lineal generalizado	170
6.10.1.	Regresión logística	170
6.10.2.	Regresión de Poisson	171
6.10.3.	Conclusiones y recomendaciones	175

7. Elicitación a partir de varios estudios: metaanálisis	176
7.1. Metaanálisis	177
7.1.1. Análisis estadístico	185
7.1.2. Modelo de efectos fijos	185
7.1.3. Modelo de efectos aleatorios	186
7.1.4. Heterogeneidad	187
7.1.5. Sesgo de publicación	188
8. Combinación de a priori y análisis preposterior	189
8.1. Combinación de a priori	189
8.2. Análisis preposterior	191
8.2.1. Distribución predictiva a priori	193
9. Apéndices	196
9.1. Programas creados para elicitación de distribuciones de probabilidad .	196
9.1.1. Elicitor	196
9.1.2. MATCH	196
9.2. Programas en <i>R</i>	197
Bibliografía	199

Índice de figuras

3.1.	<i>Plantilla para elicitar distribuciones a priori a mano alzada</i>	41
3.2.	<i>Distribución a priori a mano alzada para la edad de una persona . . .</i>	49
3.3.	<i>Función de distribución de la edad de la persona que fue elicitada a mano alzada junto con la función de distribución de una normal con media 47.7245 y varianza 9.670172</i>	51
4.1.	<i>Distribución a priori del ingreso bruto por día en miles de pesos. . .</i>	80
4.2.	<i>Funciones elicidadas a 12 sujetos (expertos) sobre la edad de un individuo</i>	82
4.3.	<i>Dendrograma de distribuciones a priori elicidadas</i>	83
5.1.	<i>Función de distribución acumulada de la estatura de los hombres en Medellín</i>	86
5.2.	<i>Distribución a priori simulada de la media de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente de nuestro conocimiento a priori igual a 10</i>	88
5.3.	<i>Distribución a priori simulada de la media de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente a nuestro conocimiento a priori igual a 30</i>	89
5.4.	<i>Distribución a priori simulada de la varianza de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente a nuestro conocimiento a priori igual a 30</i>	90
5.5.	<i>Histograma y curva normal ajustada a partir de la elicitación de los parámetros del modelo.</i>	96
5.6.	<i>Distribución elicitada paramétrica de la proporción del número de estudiantes que ejercen la prostitución como medio de supervivencia . . .</i>	103
5.7.	<i>Error absoluto cometido vs. media de la distribución de Poisson . . .</i>	108
5.8.	<i>Error cometido hasta el punto de truncamiento</i>	108
5.9.	<i>Distribución a priori para la media de la Poisson.</i>	110
5.10.	<i>Distribución a priori para la media de la poisson con modelo gamma ajustada.</i>	111
5.11.	<i>Distribución a priori sobre la estatura promedio de los medellinenses obtenida usando una distribución poblacional elicitada y una información a priori equivalente a un tamaño muestral de 20</i>	114

5.12. <i>Distribución a priori obtenida del diálogo llevado a cabo con relación al porcentaje de los médicos alcohólicos</i>	122
5.13. <i>Distribución a priori marginal simulada para β_0</i>	133
5.14. <i>Distribución a priori marginal simulada para β_1</i>	134
5.15. <i>Distribución a priori marginal simulada para σ</i>	135
6.1. <i>Intervalo de credibilidad y distribución a priori elicitada para la proporción</i>	149
6.2. <i>Distribución de la media de la longitud del pie</i>	160
6.3. <i>Distribución de la media de la anchura máxima del pie</i>	160
6.4. <i>Distribución de la media de la anchura del talón</i>	162
6.5. <i>Distribución de la media de la longitud del dedo gordo del pie</i>	162
6.6. <i>Distribución de la media del ancho del dedo gordo del pie</i>	163
6.7. <i>Distribución conjunta posterior de la longitud máxima y anchura máxima del pie</i>	164
7.1. <i>Errores estándar de la proporción estimada (transformada) para la presencia del Helicobacter pylori</i>	184
8.1. <i>Ilustración del flujo para la elicitación de un grupo de expertos</i>	191
8.2. <i>Distribución predictiva a priori</i>	195

Índice de tablas

2.1.	<i>Estimación del total para diferentes intervalos de tasas de error . . .</i>	28
3.1.	<i>Media, mediana, desviación típica y rango de la cuantificación correspondiente a 43 expresiones probabilísticas</i>	45
3.2.	<i>Escala verbal para siete expresiones probabilísticas</i>	46
3.3.	<i>Valores medios estimados para distintas expresiones</i>	47
3.4.	<i>Cuantiles estimados vía método de bisección</i>	55
3.5.	<i>Aproximaciones para la varianza, la moda, la y media de la distribución gamma con diferentes combinaciones de parámetros</i>	59
4.1.	<i>Parámetros estimados del modelo gamma para cada grupo</i>	79
5.1.	<i>Distribución de frecuencias para el número de goles marcados por los equipos locales en el torneo de fútbol profesional colombiano</i>	106
5.2.	<i>Tamaños de muestras correspondientes al nivel de seguridad del experto respecto a su información a priori</i>	112
5.3.	<i>Distribución subjetiva del cliente</i>	122
5.4.	<i>Valores elicitados por grupos de edad</i>	142
5.5.	<i>Estimados vía elicitación por grupos de edad</i>	144
6.1.	<i>Distribución de frecuencias de los tipos de sangre</i>	150
6.2.	<i>Distribución de frecuencias de los tipos de sangre</i>	153
6.3.	<i>Parámetros estimados del modelo</i>	154
6.4.	<i>Medidas de pies para 10 estudiantes</i>	167
7.1.	<i>Distribución porcentual de la prevalencia de insomnio y somnolencia en 230 estudiantes</i>	177
7.2.	<i>Tasa de éxito observada con la SAME para nueve lugares de ensayo .</i>	179
7.3.	<i>Prevalencia del Helicobacter pylori en diferentes regiones de Colombia (Bravo et al., 2003)</i>	181

Prefacio

En muchos problemas de toma de decisiones el investigador se enfrenta con situaciones en las que una o más personas conocen el problema o parte de él a un nivel en el cual sus opiniones deben ser consideradas en la solución, pero esta información no es fácil de expresar en términos de una o más distribuciones de probabilidad, la cual refleja tanto el nivel de conocimiento como el nivel de confianza en este. El problema fundamental del progreso científico, y uno fundamental en la vida diaria, es el de aprender de la experiencia. El conocimiento obtenido de esta manera es parcialmente una descripción de lo que ya hayamos observado, pero una parte consiste en la realización de inferencias de la experiencia pasada para predecir la experiencia futura (Jeffreys, 1961).

La cuantificación es un problema complejo y en muchas circunstancias se prefiere usarla marginalmente o no hacerlo. Esto, obviamente, no es posible en situaciones en las que es necesario realizar evaluaciones de riesgo, cuando el juicio de expertos en el área de interés es el único elemento que se puede utilizar; por ejemplo, cuál será el resultado de una cirugía cerebral en un paciente específico, el resultado del lanzamiento de un satélite al espacio, la inversión en una nueva acción, etc. La selección de un procedimiento de elicitación apropiado de probabilidades debe tener en cuenta tanto la estructura del problema como el tipo de experto que va a proporcionar la información (van Noortwijk et al., 1992). El problema de la cuantificación de información personal es tan importante y complejo que organismos de defensa nacional como el británico han sacado manuales sobre elicitación (DSTL, 2014).

La mayoría de las veces utilizamos procedimientos informales para obtener este tipo de información. ¿Quién no tiene una idea del tiempo promedio que tarda en llegar desde su trabajo hasta su lugar de residencia permanente? Un estudiante que acaba de presentar un examen de Cálculo tiene alguna idea de cuál va a ser la nota que su examen va a recibir por parte del docente. Lamentablemente, en la literatura no abundan los documentos en los que se trate en forma metódica este tema. A pesar de tener una literatura relativamente antigua en ciertas propuestas, estas no han sido sometidas a validación científica con experimentos controlados, con excepción del trabajo de Barrera-Causil et al. (2019). Esta es un área muy incipiente que apenas está en sus inicios a pesar de su importancia.

El objetivo de este trabajo, por lo tanto, será presentar diferentes formas de cuantificar información personal de tal forma que dispongamos de metodologías empíricas claras y que puedan ser replicadas en casos en los que sea necesario.

Introducción

Nadie niega que ante un problema debemos utilizar toda la información disponible acerca de este en particular. Esta información puede obtenerse de muy diferentes formas: publicada, generada a partir de experimentación, y a partir de personas que conozcan y que hayan estudiado el problema o problemas similares. La forma tradicional de proceder en estadística clásica es asumir un modelo matemático que depende de unos parámetros y que las observaciones reales son generadas desde este modelo. Hay muchas preguntas que se hacen y que son difíciles de responder, tales como ¿quién garantiza que el modelo matemático sí sea el que realmente genera los datos?, o ¿cómo garantizamos que las observaciones que tengamos nos permitan ver una foto bien aproximada del sistema que los generó? La incertidumbre sobre parámetros poblacionales se resume por medio de distribuciones de probabilidad y, como esto se hace a partir de los datos recogidos, entonces se desarrollan procedimientos matemáticos que nos determinan las distribuciones “muestrales”, las cuales permiten responder preguntas de tipo estadístico sobre los parámetros poblacionales. Muchas veces, se tiene información previa sobre los parámetros antes de recoger información muestral relevante para ellos, que se debe expresar en forma de distribución y es conocida como “distribución a priori, distribución personal o distribución subjetiva.” El problema está en la forma de cuantificar esta información sin generar alguna contradicción con lo que el sujeto piensa.

La escuela bayesiana en estadística ha tomado fuerza en los últimos años debido a su potencial para resolver problemas que no se pueden atacar con otros métodos y porque permite incorporar naturalmente información que es útil en la solución del problema enfrentado. La aproximación bayesiana es una herramienta fundamental en situaciones en las que la recolección de información muestral es muy difícil, por ejemplo, en tópicos de alta sensibilidad social, tales como el consumo de drogas ilícitas, o extremadamente costosos o imposibles, como sería el caso de la determinación del riesgo de falla de una nueva nave espacial o sobre la probabilidad de que haya vida inteligente en nuestra galaxia. Cressey et al. (2016) dicen:

La elicitación vía expertos se refiere a un enfoque sistemático para obtener y sintetizar juicios subjetivos de expertos sobre un tema en el que existe incertidumbre debido a datos insuficientes o cuando dichos datos son inalcanzables debido a limitaciones físicas o falta de recursos. Busca

hacer explícito y utilizable el conocimiento y la sabiduría inédita de los expertos, con base en su experiencia y conocimientos acumulados. Esto puede incluir información sobre las limitaciones, fortalezas y debilidades del conocimiento publicado y los datos disponibles.

Un problema que se ha planteado cuando se habla de la escuela bayesiana es que dos personas enfrentadas a una situación y una decisión por tomar, y asumiendo que tengan la misma información muestral, pueden llegar a dos decisiones opuestas si su información adicional es diferente. Greenland (2001) afirma que “los epidemiólogos perciben la especificación de la distribución a priori como impráctica y, además, pocos epidemiólogos emplearían métodos que no están disponibles en paquetes estadísticos líderes”.

En estadística realizamos y tratamos de responder preguntas con respecto a las características de una o varias poblaciones. En la aproximación bayesiana tenemos:

- La información sobre un parámetro (puede ser un vector) que se tiene se debe resumir en una distribución de probabilidad; esta será llamada la *distribución a priori*.
- Los parámetros son considerados variables aleatorias (esto no es aceptable en la estadística clásica). Muchas personas en bayesiana argumentan que el parámetro no es variable aleatoria y lo que es modelado es la incertidumbre sobre los posibles valores que pudiera tomar ese parámetro. Un ejemplo sencillo es la temperatura corporal que tiene un individuo en un momento determinado. Si no dispone de un termómetro, el individuo podría dar grados de creencia a diferentes valores de la temperatura. ¿Es arbitrario? Muchas personas argumentan que sí, pero podemos argumentar que muchas decisiones que tomamos se hacen por mero sentido común; por ejemplo, uno no saldría a la calle con el paraguas, si estuviera completamente seguro de que no lloverá.
- La información a priori puede provenir de:
 - Estudios previos.
 - Información subjetiva de expertos (la cuantificación de esta información es lo que llamamos *elicitación*).

Oakley (2010) habla de dos tipos de probabilidades subjetivas que describen incertidumbre epistémica, la cual se debe a falta de conocimiento, por ejemplo, la fecha de nacimiento de la pintora Débora Arango, y la incertidumbre aleatoria, que se origina por la incertidumbre del fenómeno *per se*, por ejemplo, la probabilidad de que un motociclista sufra un accidente de tránsito.

Gill y Walker (2005) dicen: “Actualmente casi no hay trabajo en las ciencias sociales que recurra a a priori elicitedas dentro de una especificación bayesiana.”

Ayyub (2001b) describe los inicios formales de la utilización de técnicas de elicitación por la corporación RAND, la fuerza aérea norteamericana y la fábrica de aviones Douglas en los años 50, con el desarrollo del método Delphi y el análisis de escenarios. Galway (2007), trabajando para la corporación RAND, hace una revisión de la elicitación de probabilidades para la determinación de análisis de costos de riesgos, un tema que día a día es más importante en las grandes corporaciones. Fauss (2008) propone métodos de elicitación en nanotecnología. Landquist et al. (2017) describen el proceso de elicitación para múltiples expertos evaluando riesgos en la industria naviera. Van der Pas et al. (2012) realizan un gran estudio de elicitación con varios expertos para estudiar accidentalidad vial. Wiser, R., Rand, J., Seel, J. et al. (2021) predicen el porcentaje de reducción del costo de la energía eólica en 2050. Esto último nos lleva a pensar en la utilidad que pueden tener las técnicas de elicitación para hacer predicciones en torno a los mercados de baterías eléctricas o a base de hidrógeno, lo cual es una incertidumbre grande para los fabricantes y compradores de automóviles hoy en día.

La estadística clásica y la información previa

La estadística tradicional (clásica y frecuentista) utiliza el conocimiento previo de varias formas:

- *Determinación de tamaños muestrales.* Usualmente, las fórmulas para la determinación de tamaños muestrales requieren estimaciones de los parámetros que se van a estimar (¿una contradicción?). Sin embargo, estas fórmulas asumen que estos valores son los correctos para determinar los tamaños muestrales que cumplan los requisitos establecidos para intervalos de confianza o para pruebas de hipótesis.
- *Para establecer hipótesis a verificar.* Estas hipótesis surgen a partir de estudios previos o teorías desarrolladas por miembros de alguna comunidad académica o científica. Por ejemplo, los economistas trabajan con series de tiempo que se modelan y extrapolan. Estas proyecciones son, en términos reales, hipótesis.
- *Para definir diseños experimentales.* En diseño experimental se utiliza para la determinación de diseños, niveles en los cuales se fijan los factores, número de réplicas, etc.

Realmente, el uso de la información no muestral en la estadística clásica es muy restringido. Sin importar qué tan buena sea la información previa, por ejemplo, estudios realizados, el estadístico clásico es como si empezara casi de nuevo. Los trabajos realizados por Pratt, Raiffa y Schlaifer, R. (1964), Raiffa (1970) en la teoría

de la decisión presentan un área donde se trabaja parcialmente con información recogida a partir de expertos.

Software para elicitación

Existen muchos programas desarrollados para fines específicos con metodologías limitadas para elicitar algunos de los parámetros más comunes como son la proporción y la media. Devilee y Knol (2011) presentan un listado de los disponibles a la fecha. Lamentablemente, la mayoría de los programas son muy limitados en las metodologías de elicitación de esos parámetros. En un apéndice, al final, presentamos un listado de programas para elicitación.

Capítulo 1

Probabilidad subjetiva o personal

El trabajo estadístico descansa en el concepto de probabilidad. La definición matemática es clara: es una función aditiva no negativa cuyo máximo valor es la unidad. El problema fundamental está en la forma como se determine esa función. Para los matemáticos puede ser relativamente simple asumir una forma cerrada, por ejemplo, el modelo gaussiano, pero en muchas ocasiones el investigador, en el contexto de teoría de decisiones el decisor, tiene en su mente ideas muy vagas de cómo puede ser este modelo. Ashby (2006) comenta:

Tres interpretaciones se le pueden dar a las distribuciones a priori: como distribuciones de frecuencia basadas quizá en datos previos, como representaciones normativas y objetivas de lo que es racional creer acerca de un parámetro o como una medida subjetiva de los que un individuo particular realmente cree.

1.1. Probabilidad

La impredecibilidad o falta de certeza en muchas de las situaciones corrientes acerca del mundo que nos rodea (tanto del futuro como del presente y el pasado) nos genera situaciones en las que, dependiendo de las consecuencias que este desconocimiento tenga en un momento determinado, queramos asignarle algún indicador. Lindley (2013) dice:

Existen algunas afirmaciones que usted sabe que son ciertas, otras que sabe que son falsas, pero con la mayoría de las afirmaciones usted no sabe si son ciertas o falsas; decimos que, para usted, estas afirmaciones son *inciertas*.

Lo primero surge del hecho que el conjunto de afirmaciones que usted conoce que son ciertas difiere de mi conjunto, ya que usted conoce cosas que yo no conozco. Igualmente, cosas que son inciertas para usted pueden

ser conocidas para mí; pero hay más aún, si tomamos una afirmación la cual los dos estamos inseguros, usted puede tener más confianza en que sea cierta que yo; diferimos en nuestro grado de incertidumbre.

Hay *dos* cálculos de probabilidad, aparentemente distintos. En uno, las probabilidades son definidas sobre frases; y en el otro, sobre subconjuntos de algún conjunto dado. El segundo, desarrollado desde la teoría de conjuntos, fue introducido en el siglo XX por Kolmogorov, aunque la teoría matemática de las probabilidades data del siglo XVII. La primera teoría asignaba probabilidades directamente a hipótesis (o tipos de frases o afirmaciones). Franklin (2001) lo dice:

La probabilidad es de dos clases. Hay una probabilidad *factual* o *estocástica* o *aleatoria*, que trata con las elaboraciones aleatorias tales como el lanzamiento de un dado o una moneda, que producen sucesiones aleatorias o sin patrones característicos. De otra parte, existe una probabilidad *lógica* o *epistémica*, o *lógica no deductiva*, que se involucra con la relación de confirmación o soporte parcial entre una proposición y otra.

La presentación de probabilidades basada en frases tiende a ser favorecida por personas interesadas primariamente en la aplicación o en el problema de la inferencia inductiva, debido a que las entidades básicas sobre las que hace referencia la probabilidad son explícitamente entidades lingüísticas, es decir, hipótesis de una forma u otra.

La diferencia entre los dos tipos de desarrollos es más bien formal: los conjuntos en la función de probabilidad de Kolmogorov son realmente hipótesis disfrazadas.

Asumamos que S es una clase dada de afirmaciones a, b, c, \dots , que pueden contener conjunciones, disyunciones y negaciones de cualquiera de las frases contenidas. S puede ser cerrado bajo las operaciones funcionales de verdad; esto es, S podría contener $a \wedge b$ (a es cierta y b es cierta, conjuntamente), $a \vee b$, $\neg a$, $\neg b$, cada que contenga a y b . Esto no siempre es el caso. Lo que se asume siempre es que S es no vacío y que al menos contiene una tautología (es una frase que es cierta, independientemente del estado del mundo. Se crean fácilmente a partir de afirmaciones del estilo ' $a \vee \neg a$ ').

Una función de probabilidad sobre S es una función que asigna números reales no negativos a las afirmaciones en S , de tal forma que a cada tautología se le asigna 1, y la suma de las probabilidades de dos afirmaciones mutuamente inconsistentes es igual a la probabilidad de su conjunción; en otras palabras, las siguientes cuatro condiciones o axiomas, son satisfechas:

1. $P(a) \geq 0$ para toda a en S
2. $P(t) = 1$ si t es una tautología

3. *Principio de aditividad:* $P(a \vee b) = P(a) + P(b)$ si a , b y $a \vee b$ están en S , y a y b son mutuamente inconsistentes (una afirmación implica la negación de la otra).

4. *Probabilidades condicionales:*

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}, \quad (1.1)$$

donde a , b y $a \wedge b$ están en S y $P(b) > 0$.

Las tres primeras condiciones son suficientes para generar esa parte del cálculo de probabilidades con las llamadas probabilidades absolutas o incondicionales.

1.2. Clasificación de las distribuciones a priori

Los estadísticos bayesianos clasifican las distribuciones a priori en dos categorías que han generado varias escuelas dentro de estos.

$$\text{Distribuciones a priori} = \begin{cases} \text{propias} \\ \text{impropias} \end{cases}$$

Distribución a priori propia. es una distribución que asigna pesos no negativos y que suman o integran uno, a todos los valores posibles del parámetro.

Así, una distribución propia satisface las condiciones de función de densidad de probabilidad. Una *distribución impropia* es la que suma o integra a un valor diferente de uno, digamos K . Si K es finito, entonces la distribución impropia induce una distribución propia normalizando la función. Si K es infinito, entonces la distribución tiene un papel de ponderación o de herramienta técnica para llegar a una distribución posterior.

$$\text{Distribuciones a priori} = \begin{cases} \text{informativas} \\ \text{no informativas} \end{cases}$$

Distribución a priori no informativa. Decimos que una distribución a priori es no informativa cuando refleja una ignorancia total o un conocimiento muy limitado sobre el parámetro de interés. Una distribución a priori no informativa puede ser propia o no. En el caso de estar interesados en un parámetro como la proporción, la distribución a priori uniforme es escogida por muchos como no informativa y, sin embargo, es propia.

También, existe el concepto de una *distribución localmente no informativa*. Por ejemplo, si estamos interesados en la estatura promedio de los hombres adultos jóvenes de Medellín-Colombia, y alguien escoge una distribución uniforme de 140 cm a 200 cm, esta sería localmente no informativa.

$$\text{Distribuciones a priori} = \begin{cases} \text{conjugadas} \\ \text{no conjugadas} \end{cases}$$

Distribución a priori conjugada. Decimos que una distribución a priori es conjugada, si al proceder a su actualización mediante la información muestral, la distribución a posteriori es igual a la a priori, excepto en los hiperparámetros. El uso de este tipo de familias se hizo inicialmente con el fin de simplificar los cálculos, a pesar de que muchos autores de los años 50 y 60 se encantaban con sus cálculos algebraicos, pero su utilidad práctica se ha mostrado en sistemas inteligentes en los que las actualizaciones de las distribuciones se automatizan directamente. La elicitación debe, por tanto, llevar a la determinación de los valores de los hiperparámetros, de tal forma que la distribución coincida con el conocimiento del experto o de la información disponible en estudios previos.

1.3. Distribuciones a priori no informativas

En muchas ocasiones sabemos nada o muy poco acerca del parámetro de interés o no queremos involucrar en nuestro estudio información previa, sino más bien dejar que sean los datos los que “hablen por ellos mismos”. En este caso, la distribución debe reflejar nuestro total desconocimiento de los valores posibles del parámetro. Esta es un área de trabajo que ha crecido enormemente. El analista asume que no existe información previa o la ignora. Pero, lamentablemente, las distribuciones no informativas no son únicas, como bien es conocido para el parámetro de una distribución de Bernoulli, que existen muchas, entre ellas la a priori no informativa de Laplace y la a priori no informativa de Jeffreys, el problema termina siendo artificial, ¿cuál es la a priori no informativa que es menos informativa? o ¿cuál tiene alguna propiedad analítica que la haga mejor que las otras?

1.3.1. Distribuciones a priori informativas

Una de las mayores dificultades en la ejecución de un análisis bayesiano concierne a la identificación de la selección y la justificación de la distribución a priori. Preguntas como las siguientes deben ser resueltas sin lugar a dudas.

- ¿Qué clase de distribución a priori debemos utilizar?
- ¿Qué tipos de datos están disponibles para seleccionar el modelo a priori?
- ¿Cómo cuantificamos la información subjetiva?
- ¿Cómo ajustamos la distribución a priori con los datos subjetivos disponibles?

1.3.2. Probabilidad personal

Horowitz (1968) define la probabilidad como:

La probabilidad no es sino un número índice, entre 0 y 1, que expresa un pensamiento del individuo sobre la posibilidad del resultado, relativo, de una experiencia. Debemos por tanto, reconocer que podemos evaluar la probabilidad, bien cuando el suceso es único o se trata de un suceso de carácter repetitivo, que pueda presentarse en varias pruebas. El hecho de que el suceso vaya a ocurrir una vez no impide que un individuo pueda formar un juicio acerca de lo probable que suceda respecto a otros posibles resultados; es decir, puede asignar probabilidades a cada uno de los posibles resultados.

Para Poirier (1988)

Para un subjetivista, la probabilidad es interpretada como un grado de creencia fundamentalmente interna de un individuo como opuesto a alguna característica del mundo externo. La probabilidad subjetiva mide una relación entre el observador y los eventos (no necesariamente ‘repetitivos’) del mundo exterior, expresando la incertidumbre personal del observador acerca de esos eventos.

1.4. Probabilidad subjetiva, apuestas y loterías

Shafer (1987) dice:

Laplace, escribiendo a comienzos del siglo diecinueve, definió la probabilidad numérica como la medida de la ‘razón que tenemos de creer’. Pero a mediados del siglo diecinueve, muchos estudiantes de la probabilidad estaban buscando por una definición más empírica. Ellos hallaron esta definición en la idea de la frecuencia y procedieron a rechazar aquellas aplicaciones de la teoría de la probabilidad que no pudieran estar basadas en las frecuencias observadas. En particular, rechazaron el método de Laplace de calcular la probabilidad de las causas, que es un caso especial de la estrategia de probabilidad condicional.

La filosofía frecuentista restringe el dominio de aplicaciones de la probabilidad numérica, y aquellos quienes querían usar probabilidades numéricas más generalmente fueron forzados a buscar una fundamentación filosófica para la estrategia de la probabilidad condicional que ajustaría la posición positivista. Tal fundamentación filosófica fue establecida en el siglo veinte por Ramsey, de Finetti y especialmente Savage¹. Estos autores concibieron la idea de que a la probabilidad subjetiva se le debe dar una interpretación de comportamiento y, por lo tanto, positivista, por lo que las probabilidades de una persona deben ser deducibles de sus elecciones. Ellos formulan postulados para lo que llaman un comportamiento racional, postulados que aseguran que las elecciones de una persona determinan probabilidades numéricas. Y argumentan que es normativo que se sigan estos postulados y, por lo tanto, normativo tener probabilidades subjetivas. Poirier (1988) afirma que los frecuentistas interpretan la probabilidad como una propiedad del mundo externo.

¹Savage ha sido uno de los pilares fundamentales en la formalización de la probabilidad subjetiva. Su utilización de hombre racional impactó tanto el mundo económico como el mundo estadístico. Sus supuestos primarios son lógicos y muchos los aceptan como una forma plausible de comportamiento. Estos pueden resumirse así: Sean a_1 y a_2 acciones que pertenecen a un espacio de acciones, c_1 y c_2 posibles consecuencias en un espacio de consecuencias. Sea Ω un espacio de resultados con eventos denotados por letras mayúsculas A, B , etc.

1. Si para cualesquiera dos actos a_1 y a_2 , tenemos que $a_1 \prec a_2$ (a_1 no es preferido a a_2 o a_2 es preferido a a_1) o $a_2 \prec a_1$ o $a_1 \sim a_2$ (a_1 no es preferido a a_2 o a_2 no es preferido a a_1 , o sea que a_1 es indiferente con respecto a a_2 y viceversa).
2. Si $\Omega = B \cup B^c$ y a_1, a_2 y a'_1, a'_2 son tales que:
 - a) en B^c , a_1 y a_2 tienen las mismas consecuencias y a'_1 y a'_2 tienen las mismas consecuencias;
 - b) en B , a_1 y a_2 tienen las mismas consecuencias y a'_1 y a'_2 tienen las mismas consecuencias;
 - c) $a_1 \sim a_2$ entonces $a'_1 \sim a'_2$.
3. Si la acción a_1 tiene consecuencias c_1 y la acción a'_1 tiene consecuencias c'_1 sobre todo el espacio de resultados (o sea sin importar el verdadero estado de la naturaleza), $a_1 \prec a_2$ implica que $c_1 \prec c_2$ para algún evento B . Contrariamente, si $c_1 \prec c_2$ entonces $a_1 \prec a_2$ dado algún evento B .
4. Cualesquiera sean los eventos sobre los que un individuo decida aceptar son independientes de la cantidad a apostar.
5. Existe al menos un par de consecuencias c_1 y c'_1 tal que $c'_1 \prec c_1$.
6. Si $a_1 \prec a_2$ y si c_1 es cualquier consecuencia, existe una partición de Ω tal que si a_1 y a_2 son modificados por cualquier punto del espacio de resultados tal que las consecuencias c_1 están asociadas con este elemento y los otros elementos permanecen iguales, la acción modificada a_1 permanece estrictamente menos preferida a la acción a_2 .

Entonces, existe una única distribución de probabilidad que concuerda con el estado mental del individuo con respecto a la incertidumbre de los eventos expresados por sus preferencias (Holloway, 1979).

Es evidente que “subjetividad” no se puede confundir con “arbitrariedad”, ya que todos los elementos para la asignación de probabilidades deben tenerse en cuenta, incluyendo el conocimiento que otros puedan asignar a las diferentes posibilidades de los mismos eventos. En el fondo, la probabilidad objetiva no es más que una mentira en términos prácticos y una verdad en un sentido teórico.

Una definición formal de probabilidad subjetiva puede darse en términos de loterías como sigue. Su probabilidad subjetiva $P(E)$ del evento E es el número $P(E)$ que hace que usted esté indiferente entre las siguientes dos loterías.

Lotería A Usted gana X con probabilidad $P(E)$.
 Usted gana Y con probabilidad $1 - P(E)$.

Lotería B Usted gana X si ocurre el evento E .
 Usted gana \$0 si no ocurre el evento E .

Aquí, X y Y son dos “premios”. La única restricción sobre X y Y es que uno debe ser preferido sobre el otro; si usted es indiferente entre X y Y , entonces usted será indiferente entre las dos loterías sin importar la escogencia de $P(E)$.

Un problema con esta forma de elicitación es que es altamente demandante tanto para el elicitador como para la persona que está proporcionando la información. Esto genera agotamiento y, a la larga, la persona elicitada termina dando información a la ligera (Cooke, 1991).

Un problema que ha sido señalado para las apuestas es que realmente se asume que las personas involucradas son “agentes racionales” (Jeffrey, 1992) cuyo fin es maximizar su propia función de utilidad que, por simplicidad, se asume lineal con respecto al dinero involucrado. Obviamente, esto no es cierto para cualquier rango de dinero, ya que no todo el mundo tiene la misma percepción sobre el valor del dinero a ciertos niveles (Kadane y Winkler, 1988). Un ejemplo es: si una lotería involucra cantidades en miles de pesos, las personas actúan de forma diferente a cuando se refieren a millones de pesos. Aun para dos personas, la misma cantidad puede tener utilidades diferentes; suponga el costo de una bolsa de leche, que puede ser muy marginal para mucha gente, pero esta cantidad puede apreciarse diferente desde el punto de vista de una madre pobre con un bebé a quien alimentar.

Una función de utilidad U es una función de valor real definida en \mathbb{R} , si tiene la siguiente propiedad: sean P_1 y P_2 dos funciones de probabilidad tal que $E_{P_1}[U]$ y $E_{P_2}[U]$ existan. Entonces, P_1 será, a lo más, tan preferido como P_2 , denotado por $P_1 \succsim P_2$, si, y solo sí, $E_{P_1}[U] \leq E_{P_2}[U]$. Para cada premio $r \in R$, el número $U(r)$ es llamada la utilidad de r . El $E_P[U]$ se conoce como la utilidad de P . Los siguientes dos resultados son consecuencias de la definición de función de utilidad:

1. Si r_1 y r_2 son dos premios en R , entonces $r_1 \tilde{\succ} r_2$, si, y solo sí, $U(r_1) \leq U(r_2)$.
2. Si \wp es el conjunto de distribuciones de probabilidad de interés, entonces las funciones de probabilidad se pueden comparar. Esto es, podemos ordenarlas.

Los siguientes supuestos son establecidos en teoría de la utilidad:

- Si P_1, P_2 y P son distribuciones en la clase de referencia \wp y si α es un número cualquiera en $(0, 1)$, entonces $P_1 \prec P_2$ si, y solo sí, $\alpha P_1 + (1 - \alpha)P \prec \alpha P_2 + (1 - \alpha)P$.
- Si P_1, P_2 y P son distribuciones en la clase de referencia \wp tal que $P_1 \prec P \prec P_2$, entonces existen números α y β en $(0, 1)$ tal que $P \prec \alpha P_2 + (1 - \alpha)P_1$ y $P \succ \beta P_2 + (1 - \beta)P_1$.

Un resultado importante que permite justificar la construcción de distribuciones a priori vía loterías es el siguiente: sean r_1, r_2 y r tres premios cualesquiera en R tal que $r_1 \prec r_2$ y $r_1 \tilde{\succ} r \tilde{\succ} r_2$. Entonces, existe un único número v ($0 \leq v \leq 1$) tal que $r \sim v r_2 + (1 - v)r_1$.

Otra alternativa similar consiste en el uso de un *contrato de referencia* (Winkler, 1972). En esta situación se asume la existencia de un premio muy atractivo que depende de la ocurrencia de un evento de interés. Por ejemplo, suponga que a usted le ofrecen un contrato que le paga un millón de pesos, si su carro es robado en menos de un año (asumamos que su carro cuesta esta cantidad). ¿Cuál es la máxima cantidad de dinero que usted estaría dispuesto a pagar por este contrato? Si usted está dispuesto a pagar cien mil pesos, entonces su probabilidad subjetiva es $100.000/1.000.000 = 0.1$. Este argumento es similar al caso de las apuestas y la idea detrás es la maximización de la utilidad del individuo. Kadane y Winkler (1988) discuten la relación entre elicitación de probabilidades y utilidades y otras alternativas.

1.5. Probabilidad personal o subjetiva

Las ideas iniciales de la probabilidad surgieron relacionadas con los juegos de azar y su interpretación es básicamente frecuentista. Esta formulación frecuentista trabaja bien en muchas situaciones, pero no en todas. Aun así, piense en que usted es invitado a participar en un juego donde el resultado depende del resultado de un dado con un tahúr: ¿creería usted que el dado de él es un dado bien calibrado? ¿Podríamos pedirle que lo lanzara n veces para estimar la probabilidad de cada resultado? ¿Qué opina usted?

Kyburg, Jr. (1966) señala que pueden considerarse tres tipos de interpretaciones de la probabilidad²:

²Good (1965) presenta la siguiente clasificación de las probabilidades:

1. *Interpretación empírico-frecuentista*. Esta es la interpretación más prevalente de la probabilidad y hace relación al comportamiento (real o hipotético) de ciertos objetos.
2. *Interpretación lógica*. Esta interpretación no es común entre los estadísticos y está más bien reservada al mundo de los lógicos. De acuerdo con esta interpretación, hay una relación lógica entre una afirmación (considerada una hipótesis) y otra afirmación (considerada evidencia) en virtud de la cual la primera tiene cierta probabilidad relativa a la segunda. Probabilidad lógica es el grado de creencia en proposiciones, que asocian un conjunto de premisas con un conjunto de conclusiones. En la probabilidad lógica, esta relación es única. Fue De Morgan quien primero definió la probabilidad en términos de “grados de creencia” (Chacko, 1991).

Bajo la influencia de Bertrand Russell, Keynes adoptó una *proposición* (en lugar de un evento) “como eso que puede llevar el atributo de la probabilidad”. Keynes dice que la probabilidad es *relación lógica* indefinible entre (1) una proposición y un cuerpo de conocimiento, (2) entre una afirmación y otra afirmación(es) que representa evidencia, una relación asociada con el *grado de creencia* racional en la proposición (Chacko, 1991). Un concepto de probabilidad lógica es empleado cuando uno dice, basado en la evidencia real, que la teoría de un universo permanentemente estable es menos probable que la teoría del big bang o que la culpabilidad de un acusado, que está probada más allá de una duda razonable, no es completamente cierta (Franklin, 2001).

3. *Interpretación subjetivista*. Esta es una versión más débil de la interpretación lógica. Es más del tipo psicológico que lógico. El grado de creencia es el concepto fundamental de la interpretación: las afirmaciones probabilísticas representan los grados de creencias de los individuos (estos no son más que individuos idealizados).

-
- *Probabilidad física* (probabilidad material o probabilidad intrínseca o propensión): es considerada una propiedad intrínseca del mundo material, independientemente de la lógica y de la mente.
 - *Probabilidad psicológica*: es el grado de creencia o intensidad de la convicción que es usada para propósitos de apuestas, para la toma de decisiones o para cualquier otro objetivo, no necesariamente después de un periodo de maduración y no necesariamente con el intento de lograr una “consistencia” con las opiniones de otro.
 - *Probabilidad lógica*: es una intensidad racional de convicción, implícita en la información dada, y tal que si una persona no está de acuerdo ella está equivocada.
 - *Probabilidad intuitiva*: incluye la probabilidad subjetiva y la lógica.
 - *Probabilidad tautológica*: surge como un objeto matemático puro.

1.5.1. Grados de creencia

Ya que las probabilidades subjetivas representan grados de creencia, un estado mental, es necesaria una técnica para extraer el estado mental y convertir los sentimientos a una forma utilizable. El método particular usado puede afectar la manera en que el sujeto mira el problema que tiene en sus manos; por lo tanto, diferentes técnicas pueden producir diferentes respuestas. Uno puede utilizar técnicas de autoelicitación, o sea, el sujeto elicit sus propias respuestas o uno puede usar un entrevistador entrenado que cuestione al sujeto. De cualquier forma, se requiere un método de elicitación. Dos tipos diferentes de expresiones han sido usados en distintas ocasiones. El sujeto puede expresar sus creencias en términos de probabilidades o en términos del valor o de la escala de la variable independiente. El método particular de elicitación que debe ser seleccionado es una función de su costo, del sujeto, o de la forma del conocimiento del sujeto.

Existen teorías contradictorias sobre si la inferencia humana sigue el teorema de Bayes. Gigerenzer y Hoffrage (1995) discuten los resultados de investigadores que concluyen que el hombre adquiere información y modifica sus creencias en un modo bayesiano y otros investigadores que concluyen, a partir de sus propios experimentos, que el ser humano modifica sus creencias, pero no sigue ningún modelo. Un problema que se puede apreciar a partir de algunas aproximaciones es asumir que el cerebro y sus operaciones se pueden aceptar en cierta forma de tipo mecanicista (aunque no lo afirmen tan crudamente), reflejando el mundo newtoniano. Estas sobresimplificaciones, que en la parte aplicada a veces se hacen necesarias para obtener algunos resultados, pueden llegar a ser frustrantes en la parte experimental. Siempre debemos estar alerta a los problemas que se tienen cuando se realiza una cuantificación del conocimiento de un individuo acerca de un problema, ya que el hecho de obtener una distribución subjetiva pudiera afectar el mismo conocimiento y los sentimientos con respecto a este.

1.6. El sujeto racional

La expresión *sujeto racional* hace referencia a un sujeto ideal que sigue un conjunto de postulados en forma siempre coherente. Este sujeto, en economía, siempre busca maximizar su función de utilidad.

1.6.1. Supuestos de coherencia

Asumimos que la persona que asigna las probabilidades subjetivas mediante interrogación directa, posee las siguientes condiciones ideales:

1. Nunca viola el supuesto de coherencia.
2. Entiende perfectamente los métodos utilizados para la asignación de probabilidades. Esto es, entiende las alternativas que se le presentan y las implicaciones de cada alternativa.
3. Tiene una función de utilidad que es lineal con respecto al dinero en el rango relevante. Además, escoge sus respuestas de tal forma que maximiza su utilidad esperada. Esta condición ha sido objeto de revisiones y críticas, ya que se ha mostrado empíricamente que los sujetos a menudo violan este supuesto (Tversky y Kahneman, 1992).

Lo anterior está ligado a lo que se conoce como *consistencia y coherencia interna*. Esto hace referencia a lo bien que un experto ajusta sus probabilidades internas, de tal forma que estas concuerden con la definición formal de probabilidad y que el sujeto no va a entrar en contradicciones técnicas. Este es un tema que ha generado mucha discusión y es considerado un punto débil en cualquier proceso de elicitación de probabilidades subjetivas. Creemos que un proceso de elicitación puede ser comparado al de la medición del nivel de inteligencia de una persona, donde tanto la conceptualización de lo que es ser inteligente como el instrumento de medición afectan el resultado final. En un proceso de estos es posible encontrar resultados que pueden ser contradictorios y esto se debe a los actores que entran en el proceso (experto y elicitor), y, además, el instrumento de elicitación. ¿Será malo aquel sujeto que aparentemente se contradiga en algunas opiniones cuantificadas en términos probabilísticos? Pensamos que esto es parte del proceso y que muestra que los procedimientos deben ser revaluados en forma dinámica durante el mismo.

Este sujeto racional en la realidad no existe y esto, desde el punto de vista teórico, genera dificultades, pero no mayores que desde el punto de vista aplicado, donde nos enfrentamos con sujetos que pueden generar situaciones inesperadas. Podemos resumir lo anterior de la siguiente forma:

- Ya que las probabilidades subjetivas representan grados de creencia, un estado mental, es necesaria una técnica para extraer el estado mental y convertir los sentimientos a una forma utilizable.
- El método particular usado puede afectar la manera en que el sujeto mira el problema que tiene en sus manos; por lo tanto, diferentes técnicas pueden producir distintas respuestas.

- Uno puede utilizar técnicas de autoelicitación, o sea, el sujeto elicitó sus propias respuestas o uno puede usar un entrevistador entrenado que cuestione al sujeto. De cualquier forma se requiere un método de elicitación.
- Dos tipos de expresiones han sido usadas en diferentes ocasiones: el sujeto puede expresar sus creencias en términos de probabilidades o en términos del valor o de la escala de la variable independiente.
- El método particular de elicitación que debe ser seleccionado es una función de su costo, del sujeto o de la forma del conocimiento de este.

Li y Krantz (2005) han hallado que los individuos a los que se les elicitó su distribución de probabilidad subjetiva acerca de un parámetro, tienden a tener mayor confianza que la que debieran; por ejemplo, cuando construyen intervalos de probabilidad del 80 %, estos corresponden realmente a un nivel del 50 %. Sin embargo, y pese a estos problemas, ellos hacen énfasis en casos en los que recoger información muestral es muy costoso o imposible, entonces, se deben elaborar instrumentos de elicitación que sean fiables. También, el uso de diferentes procedimientos puede ayudar en el refinamiento de la elicitación. Brenner et al. (1996) discuten el problema del exceso de confianza en juicios sobre probabilidades.

- *Falacia de conjunción*: usualmente no es aplicable a la elicitación de expertos, pero consiste en asignarle probabilidades mayores a un subconjunto de un evento que las que se le asignan a este mismo evento. Como ejemplo consideremos un experimento que se realizó por Tversky y Kahnemann en 1982 (Burgman et al., 2007), conocido como “el problema de Linda”, donde ellos le presentan la siguiente información a un sujeto: Linda es una mujer de 31 años, soltera y segura de sí misma; estudió filosofía y como estudiante estuvo profundamente involucrada en problemas de discriminación. Se le pide al sujeto que escoja una de las dos siguientes posibilidades como la más probable:

1. Linda trabaja como cajera en un banco.
2. Linda trabaja como cajera en un banco y es activista en un movimiento feminista.

La mayoría de la gente escoge la segunda opción como más probable que la primera, siendo la segunda un subconjunto de la primera y, por lo tanto, siendo su probabilidad a lo menos igual a la segunda.

- *Sesgo de prevista*: ocurre si el asesor ha visto los datos muestrales previamente, por lo tanto, su información elicitada ya ha sido actualizada con base en estos datos.

Butterworth (1988) describe una situación experimental en la que los expertos producen resultados tan malos como los de personas no expertas en la determinación

de calificaciones obtenidas por estudiantes. Se especula en este estudio que el exceso de confianza pudo ser una de las causas para esto. “Una elicitación es bien hecha si la distribución es obtenida de tal forma que representa de manera precisa el conocimiento del experto, sin considerar qué tan bueno sea este conocimiento” (Garthwaite, Kadane, y O’Hagan, 2004). Gigerenzer y García-Retamero (2017) hablan de la *lamentación de Cassandra*, situación en la cual el experto se niega a aprender de la experiencia, muy común entre la gente que sufre eventos indeseados y, aun teniendo mucha información sobre la ocurrencia, se niegan a cambiar sus probabilidades.

Kynn (2008) reconoce que en un proceso de elicitación no solo es importante lo que se pregunta, sino cómo se pregunta. Más allá de medidas de “bondad” de un procedimiento de elicitación, el criterio primario en la selección de un proceso de elicitación es su practicidad. Si el experto puede responder las preguntas y se siente confortable, al final, si su opinión ha sido de alguna manera capturada y, asumiendo que el método cumpla con los criterios matemáticos básicos de coherencia, y que involucre algún procedimiento para probar su confiabilidad, es un buen método.

Capítulo 2

Elicitación

Uno de los problemas más difíciles para el uso de la estadística bayesiana ha sido el de la elicitación¹ de probabilidades subjetivas de información personal sobre una situación particular. “La elicitación es el proceso de formular el conocimiento y creencias de una persona acerca de una o más cantidades inciertas en una distribución de probabilidad (conjunta) para estas cantidades” (Garthwaite, Kadane, y O’Hagan, 2004). Sin embargo, Stangl y Berry (1998) anotan que la elicitación es uno de los procesos más desafiantes dentro de la estadística bayesiana, ya que:

Mientras los investigadores están más que dispuestos a usar información a priori para determinar qué estudiar, estimar los tamaños muestrales requeridos y decidir si un estudio previo es convincente, ellos son algunas veces reacios para convertir formalmente sus creencias en una distribución a priori. Los expertos rara vez concuerdan. Y, la gente tiene dificultades cuantificando sus creencias en términos probabilísticos. Un problema especial es que los que asignan probabilidades usualmente subestiman su incertidumbre.

Hahn (2006) presenta cuatro principios básicos en el proceso de elicitación de distribuciones informativas. Estos son:

- **Principio 1.** *Es deseable para las metodologías de elicitación producir distribuciones que tengan formas flexibles.* Los procedimientos de elicitación deben darle oportunidad al experto de expresar sus incertidumbres en formas no necesariamente estándares; por ejemplo, la unimodalidad en muchos procesos se presume como una condición obligatoria.

Este principio es uno de los más difíciles de sostener, ya que usualmente los investigadores pretenden obtener una única y bien definida distribución de pro-

¹La palabra *elicitación* no existe en castellano. No obstante, vamos a usar este anglicismo, ya que su traducción no refleja el espíritu del procedimiento y, además, no hay un término universal en español que nos dé la idea precisa de este proceso.

babilidad que represente sus juicios. En este documento se propone una metodología que nos permite construir conjuntos de distribuciones que representan la información del sujeto, o que podríamos llamar distribución *fuzzy* o difusa.

- **Principio 2.** *Es deseable minimizar las demandas cognitivas que una metodología de elicitación pone en el experto.* Un experto en una específica puede ser un desconocedor de estadística y sus metodologías y ningún curso acelerado lo volverá una persona hábil en esta disciplina. Esta realmente es una de las tareas fundamentales del grupo a cargo de realizar la elicitación, que debe contar con la ayuda de psicólogos, lingüistas, quienes deben ayudar a elaborar preguntas y cuestionarios que cumplan con este principio.
- **Principio 3.** *Es deseable minimizar las demandas que una metodología de elicitación pone al estadístico.* Muchas aplicaciones y experiencias en elicitación presumen que un experto entiende conceptos estadísticos que se asocian a poblaciones, por ejemplo, percentiles extremos y en caso de niveles de seguridad o confianza en el conocimiento. Estas demandas pueden ser extremas y obligar al experto a dar respuestas poco fiables.
- **Principio 4.** *Asumiendo que todo lo demás se mantiene igual, las metodologías para la elicitación de a priori que pueda ser fácilmente aplicable a una amplia variedad de modelos o escenarios puede ser más deseable.* Es labor del estadístico desarrollar procedimientos de elicitación que sean aplicables a situaciones diferentes, por ejemplo, determinación de parámetros naturales (como son las medidas de localización) y también a parámetros que no son intuitivos como las covarianzas.

Una característica distintiva de la estadística bayesiana es que tiene en cuenta, de forma explícita, la información previa y se involucra en el análisis en forma de distribución de probabilidad, llamada distribución a priori o distribución subjetiva. La teoría clásica la considera, al menos en alguna de sus características, ok promedio o valor modal, básicamente para determinar tamaños muestrales y el diseño de experimentos y, a veces, como forma de crítica de los resultados hallados².

²La gran mayoría de bayesianos no son muy amigos de los procesos de elicitación y prefieren trabajar con una clase de distribuciones a priori llamadas *no informativas*, que se supone no contienen información sobre los parámetros de interés, y prefieren que sean los datos muestrales (asumiendo que las muestras realmente son aleatorias, lo que en muchos casos es cuestionable) los que permitan hacer todo el trabajo inferencial vía la distribución posterior. Estos bayesianos tibios han sido muy afectados por las críticas de los colegas clásicos y por otros científicos que manejan un concepto muy ortodoxo de la noción de *objetividad*. Muchos bayesianos trabajan en esta clase de distribuciones más por placer académico que por utilidad práctica. En las últimas tres décadas los bayesianos se han apropiado como un invento propio los métodos computacionales conocidos como MCMC (Monte Carlo Markov Chain) que han sido un reconocimiento al poder de la simulación en la solución de problemas aplicados complejos, pero esta metodología fue desarrollada por físicos en los años 50.

La expresión de la información previa en forma cuantitativa puede ser un proceso complejo y delicado, aunque se han hallado soluciones que pueden llegar a parecer extrañas, como el uso de lo que se conoce como distribuciones no informativas, pero que se utilizan extensamente en el trabajo bayesiano aplicado.

El proceso de elicitación es un paso fundamental en el análisis bayesiano, pero en muchos trabajos académicos se reduce a una frase que sobresimplifica la complejidad de este proceso.

2.1. Fuentes de información para la construcción de a priori informativas

Bormane et al. (2016) clasifican las técnicas de elicitación en cuatro grupos:

1. Técnicas tradicionales.
 - Entrevistas
 - Cuestionarios
 - Recolección de datos existentes
2. Técnicas colaborativas. Se usan mucho en ciertos entornos sociales, tales como los negocios.
 - Grupos focales
 - Tormentas de ideas
 - Prototipos
 - Reuniones de trabajo
 - *Storyboarding*
 - Modelos
 - Uso de casos y escenarios
3. Técnicas cognitivas.
 - Análisis de documentos
 - Análisis de protocolos
 - Técnicas de escalera y rejilla de repositorios
4. Técnicas observacionales. Se basan en la observación y estudios sociales.

Dos fuentes tradicionales para la construcción de la distribución a priori se explican a continuación:

2.1.1. Estudios previos similares

La utilización de estudios previos sobre unos pocos parámetros específicos ha dado origen a un área conocida como *metaanálisis*, la cual puede trabajarse desde el punto clásico y bayesiano (Sinha et al., 2011). Un problema cuando se trabaja con datos sacados de publicaciones y no con los estudios originales, es el que se conoce como *sesgo de publicación*, que hace referencia a la divulgación, por parte de los editores o autores, de solo aquellos resultados que son significativos. De Santis (2006) describe una forma de desarrollar distribuciones a priori, conocidas como a priori de tipo potencia, la cual se obtiene al combinar una a priori con una verosimilitud calculada a partir de información muestral histórica pero ponderada por un factor que refleja de alguna forma la relevancia de esta información. Si \mathbf{z}_{n_0} denota los datos históricos y la verosimilitud del parámetro $\boldsymbol{\theta}$ es $L(\boldsymbol{\theta} | \mathbf{z}_{n_0})$ y si $\xi_0(\boldsymbol{\theta})$ denota la a priori (podríamos decir una pre a priori), entonces la a priori geométrica se define como:

$$\xi^P(\boldsymbol{\theta} | \mathbf{z}_{n_0}) \propto \xi_0(\boldsymbol{\theta}) [L(\boldsymbol{\theta} | \mathbf{z}_{n_0})]^{a_0}, \quad (2.1)$$

donde $0 < a_0 < 1$ es llamado un parámetro de precisión que valora la calidad de la información muestral previa. Como puede ser difícil especificar un valor preciso para a_0 , se puede considerar una a priori sobre este parámetro, digamos $\xi(a_0 | \gamma_0)$ y, por lo tanto, el resultado final será:

$$\int \xi^P(\boldsymbol{\theta} | \mathbf{z}_{n_0}, a_0) \xi(a_0 | \gamma_0) da_0 \quad (2.2)$$

2.1.2. Opinión de expertos

La utilización de expertos es casi obligatoria en situaciones nuevas en las que experimentar puede ser muy costoso o imposible, por ejemplo, en la aplicación de políticas macroeconómicas. Se debe estimular al experto para cuantificar su conocimiento personal a mediante distribuciones de probabilidad. Se ha utilizado en problemas complejos donde los métodos tradicionales no son aplicables (Aaron et al., 1995; Truong y Heuvelink, 2013; Nemet et al., 2013; Wilcox et al., 2016). Desafortunadamente, esta es una tarea compleja que ha sido descuidada por muchos analistas que prefieren trabajar de forma sencilla con las distribuciones a priori no informativas. Biedermann et al. (2107) dicen: “En esencia, ¿una probabilidad expresa la incertidumbre de un razonador sobre algo? ¿Por ejemplo, un estado de naturaleza del pasado, presente o futuro? Eso no es completamente conocido por esta persona”. O’Hagan y Oakley (2004) y Kadane y Wolfson (1998) discuten las dificultades del proceso de elicitación. La proporción es el parámetro que ha recibido la mayor atención en el desarrollo de métodos de elicitación no paramétricos y paramétricos (Winkler, 1967a, 1967b; Raiffa, 1970; Chaloner y Duncan, 1983; Gavasakar y Umesh, 1988). Tversky (1974) discute los problemas que surgen en cualquier proceso de elicitación de probabilidades subjetivas. Gzyl et al. (2017) consideran el problema de los expertos que no proporcionan estimaciones puntuales de los parámetros de interés,

sino intervalos de incertidumbre. DeWispelare et al. (1995) describen en detalle los pasos para ejecutar un proceso de elicitación.

El uso de conocimiento de expertos es aceptado en circunstancias en las que datos empíricos son escasos o no disponibles. Un término como *conocimiento experto* es lo que las personas calificadas conocen como resultado de sus conocimientos y estudios técnicos a cualquier nivel (tecnológico, académico o autoestudio), de prácticas en campo, entrenamiento y experiencia. Puede incluir hechos o pruebas registrados, inferencias hechas por el experto sobre la base de hechos concretos en respuesta a situaciones nuevas o indocumentadas, y la integración de diferentes fuentes en modelos conceptuales para abordar problemas del sistema (McBride y Burgman, 2012). El propósito de la elicitación es facilitar que el experto nos diga lo que él cree en términos probabilísticos, mientras se le reduzca la necesidad de que sepa mucho de probabilidad o de estadística.

Kadane y Wolfson (1998), recogiendo la opinión de varios autores, concluyen:

- La opinión de los expertos es la más valiosa para elicitar. Citando a Leonard (1998): “Considero la elicitación de las creencias a priori como un método de recolección de datos de los expertos, lo cual es mucho más completo que solo coleccionar conjuntos finitos de datos”.
- A los expertos se les debe pedir que expresen su opinión sobre cantidades observables, condicionadas solo en covariables (que también son observables), o en otras cantidades observables.
- A los expertos no se les deben pedir estimaciones de los momentos de una distribución (excepto posiblemente el primer momento); a ellos se les debe pedir que determinen cuantiles o probabilidades de la distribución predictiva.
- Hay que proporcionarles retroalimentación frecuentemente durante el proceso de elicitación.
- A los expertos se les debe pedir que den asignaciones, tanto incondicionalmente como condicionalmente, basados en datos observados hipotéticos.

Muchas veces los científicos se sienten apabullados por la palabra “objetividad”, que en muchas situaciones significa “subjetividad con conocimiento”. Adicionalmente, Poirier (1998) afirma que entre los econométricos las creencias a priori aparecen en todas las investigaciones en la forma de experticia que los profesores aplican en sus investigaciones empíricas. Kausar et al. (2010) discuten las ventajas de utilizar expertos y cuestionarios en los procesos de elicitación. Un aspecto relevante es el económico. Aunque ellos discuten que la elicitación vía encuesta es fácil de implementar, esto no implica que su realización lo sea.

2.2. Elicitación a partir de expertos

El Diccionario de la Real Academia de la Lengua nos dice de un experto: Dicho de una persona especializada o con grandes conocimientos en una materia. Una tendencia es pensar en los expertos como profesionales muy conocedores de un asunto de interés. Los expertos pueden realmente ser personas del común que saben mucho de algo, por ejemplo, su propia edad, sus hábitos, su peso, etc.

2.2.1. Tipos de expertos

Un experto en un área específica no es lo mismo que lo sea en probabilidad y estadística. Kausar et al. (2010) discuten la dificultad que puede surgir al querer determinar quién es un experto. ‘La respuesta simple a la pregunta ‘¿quién es el experto?’ es que el experto es la persona cuyo conocimiento queremos elicitar; el término ‘experto’ no necesariamente significa algo más que esto’ (Garthwaite, Kadane y O’Hagan, 2004). Burgman et al. (2007) definen un experto como a aquella persona que conoce el asunto de interés y es, a la vez, capaz de comunicar su conocimiento. Muchas veces se tiende a ser demasiado optimista con la capacidad que tiene el experto de expresar su conocimiento. Existen múltiples problemas en este proceso: sesgos debidos al lenguaje, a las experiencias que vivió el experto, formas en que el cerebro registra y filtra información del mundo externo y cómo la experiencia del sujeto las reinterpreta. Se acostumbra clasificar a los expertos en tres categorías: los generalistas, los expertos en un área específica y los normativos (Slotte, Luijs y Knol, 2008).

Expertos generalistas: conocen ampliamente una disciplina y entienden los problemas técnicos de esta. En un proceso de elicitación pueden ayudar a delimitar el alcance de esta, pueden dividir el problema en subproblemas y determinar los parámetros y su importancia relativa para ser elicitados.

Expertos específicos: usualmente están asociados con investigadores altamente especializados, quienes tienen a la mano la información más relevante y actualizada. En la práctica, este tipo de experto será el personaje fundamental en el proceso de elicitación.

Expertos normativos: tienen algún tipo de conocimiento en probabilidades y pueden ayudar en la elaboración de los modelos definitivos. En investigaciones pequeñas con recursos limitados, usualmente recurrimos a expertos generales, ya que son los más disponibles.

2.2.2. El proceso de elicitación

La captura de la información que un experto posee es una de las tareas que más dificultades presenta. Vamos a presentar algunas técnicas para la captura de esta información, teniendo siempre en mente que es un proceso extremadamente delicado y frágil, que nos puede llevar a equivocaciones.

En el proceso de elicitación tenemos pasos muy claros:

1. *Seleccionar y motivar al experto*: a pesar de que la palabra experto se utiliza permanentemente para hacer referencia a personas que conocen en profundidad un cierto tópico, es posible que no conozca lo suficiente un asunto particular o que sus prejuicios sesguen su información. Buscar expertos en ciertas áreas puede ser muy complejo, pero en otras el experto no se reconoce a sí mismo como uno, algo común en diagnósticos médicos en los que el paciente es el que mejor se conoce en ciertas circunstancias.
2. *Entrenar al experto*: los procesos de elicitación con expertos requieren preparación a varios niveles, entre ellos la preparación de estos. Esta preparación incluye cartas de invitación a participar en el proceso, compartir información sobre el problema, simulacros, etc. Un proceso de elicitación no se parece en nada a un examen de conocimientos como los que se realizan para aspirar a un cargo o al sistema educativo.
3. *Estructurar las preguntas*: elicitación es un proceso de preguntar y responder. ¿Qué es una buena pregunta? Una buena pregunta es aquella que produce respuestas que son confiables y válidas de algún parámetro de interés. La confiabilidad se toma en el sentido sicométrico: las respuestas son consistentes. Esto es, cada vez que hacemos la pregunta, obtenemos una respuesta similar. La validez hace referencia a que las repuestas correspondan a algún valor hipotético que se quiera medir.
4. Elicitar y documentar los juicios del experto.
5. Verificar los resultados.

Delavande et al. (2011) estudian el efecto de cambios en las preguntas sobre las respuestas dadas por los expertos. Gigerinzer et al. (1991) discuten problemas con el exceso de confianza que pueden manifestar los expertos en un proceso de elicitación.

Spetzler y Stael von Holstein (Chesley, 1975) presentan una lista de asuntos a verificar por parte del entrevistador. Ellos sugieren que el proceso de elicitación puede dividirse en cinco etapas:

1. Motivación del sujeto mediante una compenetración e investigando los sesgos del sujeto.

2. Estructuración de la cantidad desconocida definiéndola claramente.
3. Acción para que el sujeto piense fundamentalmente acerca del problema y que evite sus sesgos.
4. Codificación de sus juicios.
5. Verificación de las respuestas, chequeando la consistencia y mirando si el sujeto cree en sus resultados. Ellos sugieren que la tendencia inicial de los sujetos es especificar distribuciones que son muy cerradas, así que se les puede ayudar a comenzar con los extremos de la situación en lugar de estimación del centro.

2.2.3. Entrevista personal

Una de las formas más populares de elicitar información de expertos es mediante la entrevista personal cara a cara (experto y entrevistador presentes). Esto ha sido modificado un poco usando teléfono³ y los chats y videoconferencias vía internet. El elicitor, de alguna manera, dirige la entrevista y condiciona las direcciones que puede tomar. El experto puede sentirse cohibido con el elicitor, en especial si este se muestra dominante. A veces, el experto prefiere no exponer sus debilidades (conocimiento débil o creencias o posturas) y no entregar las respuestas adecuadas. La comodidad y el ambiente de confianza brindados al experto son fundamentales a la hora del proceso. El elicitor debe preparar de antemano unos formularios que le brinden la posibilidad de guiar el proceso sin necesidad de atar al experto a un cuestionario rígido que lo obligue a dar respuesta a situaciones que no le permitan mostrar toda su experticia en el tema. Idealmente, la sesión inicial debería ser de contextualización en la cual al experto se le presenten los elementos del problema de interés, tratando de entregarle pistas que le permitan recordar y poner en perspectiva para la elicitación. Una elicitación cara a cara puede dividirse en tres o más sesiones. Al experto se le puede entregar material (cuestionarios y documentos) de tal forma que pueda revisarlos en su espacio y con la tranquilidad suficiente para que vaya armando las respuestas que concuerden con su conocimiento.

La grabación de las entrevistas es obligatoria tanto para detectar debilidades en el proceso (preguntas mal elaboradas, temas con dificultades, temas inacabados), de como para determinar el uso de la información obtenida de estas. Además, sirve de material de soporte a la investigación.

La entrevista vía telefónica en procesos de elicitación no es recomendada, ya que es de larga duración y el experto puede desentenderse del tema de esta con facilidad.

³El uso de la entrevista telefónica en investigación por muestreo ha sido estudiado desde diferentes puntos de vista. El libro editado por Groves y sus colegas, *Telephone Survey Methodology* en 1988, presenta estudios detallados de algunos detalles de esta metodología. Desafortunadamente, no hay estudios para entrevistas a profundidad.

La parte administrativa del manejo de expertos es importante, ya que ellos deben tener claros aspectos fundamentales del proceso. Variables tales como duración del proceso, requisitos de tiempo, honorarios y, compromisos deben estar determinadas previamente a la elicitación. Es posible que sea necesario elaborar un contrato en el que se determinen claramente las responsabilidades y compromisos de las partes involucradas en el proceso.

Confiabilidad es el término usado para describir la estabilidad del proceso de medición. Esto es, si el proceso se repitiera bajo las mismas condiciones, se esperaría que se presentaran cambios grandes en las evaluaciones que se obtienen. Idealmente, el proceso de elicitación debería ser aplicado al mismo experto en diferentes ocasiones para determinar si se presentan variaciones grandes en el proceso. Esto no es posible y puede simplemente, dentro de un solo proceso, abordarse cada tema crítico desde varias alternativas. Esto también se mezcla con un tema conocido como validez de contenido, en el que uno quiere estar seguro de que está obteniendo las respuestas adecuadas mediante las preguntas elaboradas.

2.2.4. Problemas con los expertos

Muchos investigadores han estudiado los problemas asociados con la cuantificación de probabilidades o incertidumbre en el ámbito personal. Swieringa et al. (1976) discuten los sesgos que pueden enfrentar los expertos en procesos de elicitación.

Se han propuesto varias técnicas para ayudar a los expertos a obtener su conocimiento personal en términos de distribuciones de probabilidad. Entre ellos tenemos el método de apuestas y el método de lotería (Holloway, 1979; Harrison et al., 2014; Harrison et al., 2015). Estos dos métodos requieren una suposición muy fuerte como la linealidad de la función de utilidad personal (DeGroot, 1970).

Ayyub (2001b) define el proceso de elicitación de la opinión de un experto como un proceso formal y heurístico para obtener información o respuestas a preguntas específicas acerca de ciertas cantidades. No se recomienda considerar esta información de confiabilidad absolutamente rigurosa como lo serían los métodos analíticos. Uno puede cuestionar afirmaciones como esta, ya que por el hecho de que un resultado analítico haya sido desarrollado formalmente, mediante una matemática avanzada, existen bajo este armatoste unos supuestos que en realidad terminan siendo subjetivos.

Cuando existen estudios previos y similares (esto en el fondo también es subjetivo), podemos usar las distribuciones muestrales producidas como las a priori para nuestro caso en cuestión. Una pregunta que siempre está en el aire es la siguiente: ¿Debemos usar la distribución muestral sin ajustarla y creer que se acomoda a nuestro caso específico? Ahora, en el caso de que no existan estudios o información

cuantitativa previa y solo tengamos expertos, ¿cómo extraemos su información? ¿Cómo validamos la calidad de esta información? ¿Cómo combinamos la información de diferentes expertos? ¿O la información de un experto tomada en diferentes puntos del tiempo y que, debido al mismo proceso, ha ido cambiando? Esta es un área que permanece en un estado incipiente de desarrollo y, por lo tanto, los procedimientos que presentamos hay que tomarlos con mucha reserva y aplicarlos cuidadosamente.

“Una elicitación exitosa representa en forma confiable la opinión de la persona que está siendo elicitada. No es necesariamente ‘cierta’ en algún sentido objetivista y no puede ser juzgada de esa forma. Vemos la elicitación como simplemente una parte del proceso de modelación estadística” (Garthwaite, Kadane, y O’Hagan, 2004). Este problema puede mejor replantearse de la siguiente manera: ¿Qué tan experto es el experto? La respuesta a esta pregunta puede venir de varias fuentes: de una autovaloración del experto, de una revisión de su hoja de vida o de valoraciones de otros expertos. Es por esto mismo que el concepto de calibración puede ser discutible en este contexto (Low Choy, O’Leary y Mengersen, 2009).

En muchas situaciones es de importancia cuantificar la información subjetiva que, sobre la posibilidad de la ocurrencia de un evento, sientan uno o varios individuos. La idea es entonces “desarrollar métodos que ayuden a la persona que asigna las probabilidades, las asigne de acuerdo con sus juicios” (Winkler, 1967a, 1967b). Esta tarea es difícil y puede llegar a ser frustrante.

La literatura en psicología sugiere que la gente es propensa a ciertas heurísticas y sesgos en cómo ellos responden a situaciones que involucran incertidumbre. Como un resultado, algunas formas de hacer preguntas acerca de cantidades inciertas son preferibles a otras, y parecen ser más confiables (Garthwaite, Kadane, y O’Hagan, 2004).

El estudio de las probabilidades subjetivas apunta a dos asuntos simultáneamente: el normativo y el descriptivo. La parte normativa está relacionada con el asunto de la racionalidad y la lógica. La parte descriptiva está involucrada con las creencias, sentimientos, actitudes y preferencias de las personas y no con lo que debieran ser. Por lo tanto, en esta asignación de probabilidades personales a eventos hay dos requisitos:

1. La asignación debe obedecer ciertos postulados de coherencia.
2. Las asignaciones deben corresponder a los juicios de la persona.

Es fácil chequear el primer punto, pero en el segundo una probabilidad asignada a un evento es el resultado de la interacción de numerosas creencias o juicios que solo existen en la mente de la persona. Y hay que tener en cuenta que esta asignación se realiza en un punto del tiempo y que la misma persona puede revisarlas a la luz de nueva información. Uno, definitivamente, no puede probar que las probabilidades asignadas corresponden realmente a los juicios de la persona.

Kenney y Von Winterfeldt (1991) esquematizan los pasos en un proceso de elicitación en siete partes. Discutimos cada una de estas, pero no siempre se requiere entrar en el nivel de detalle que debatimos.

1. *Identificación y selección de los expertos.* Los expertos se pueden clasificar en tres grupos (Slottje Sluijs y Knol, 2008): los especialistas, quienes conocen el problema en profundidad. Los analistas o normativos son expertos que tienen elementos para manejar el lenguaje probabilístico y son capaces de traducir el conocimiento del especialista en forma de modelos probabilísticos. También se pueden considerar otros expertos que son llamados generalistas, quienes poseen una visión amplia de todos los asuntos tratados en el proyecto y que, además, tienen la capacidad de comunicarse con los grupos de trabajo, especialistas y técnicos. Morgan (2014) discute sobre situaciones en las cuales un experto en un área puede que no tenga bases para realizar juicios predictivos informados.

Los expertos deben tener claro cuáles son sus tareas en el proceso y cuánto tiempo estarán participando en él, y estar de acuerdo con la retribución económica que recibirán. Ayyub (2001b) presenta una lista de lo que se les debe enviar a los expertos antes del proceso de elicitación:

- a) Una descripción del objetivo del estudio.
 - b) Una lista de los expertos, observadores, integradores, facilitadores, líder del estudio, patrocinadores y elementos bibliográficos pertinentes al problema a resolver.
 - c) Una descripción de los lugares y facilidades logísticas que se tendrán durante el proceso.
 - d) La terminología básica. Las definiciones, las cuales deben incluir elementos de probabilidad, valores medios, medianas, modas, etc.
 - e) Consecuencias de una estimación insatisfactoria.
 - f) Una descripción del proceso de elicitación.
 - g) Un ejemplo relacionado de elicitación y sus resultados, si está disponible.
 - h) Métodos de agregación de los resultados de la elicitación.
 - i) Una descripción de los asuntos y parámetros en forma de preguntas con toda la información de contextualización posible.
2. *Presentación y discusión de los eventos y cantidades inciertas.* Una vez el tema y todas las cantidades de incertidumbre a ser elicítadas han sido claramente definidos, y cuando los expertos han sido seleccionados, deberá hacerse una reunión donde se haga una presentación y socialización del problema y de los expertos.

3. *Entrenamiento en juicios probabilísticos.* Una vez los expertos se han enfrentado a los temas de trabajo, deberán recibir algún tipo de entrenamiento con los conceptos y técnicas de elicitación. Conceptos tales como variabilidad o ruido de un sistema deben ser entendidos relativamente bien, así como la variabilidad en general. El experto debe también tener alguna discusión sobre los posibles sesgos que, por su parte, pueden intervenir en el proceso y sus formas de manejarlos. Biedermann et al. (2013) hablan del *scoring* de afirmaciones probabilísticas como un método para uno decir lo que quiere significar en términos probabilísticos.
4. *Análisis y recolección de los datos.* Una vez que los expertos han sido expuestos al tema a ser estudiado, se les debe permitir un tiempo para que tengan la oportunidad de madurar su conocimiento. Se puede entregar documentación, se puede permitir que los expertos interactúen entre ellos de tal forma que compartan experiencias, etc.
5. *Presentación y discusión de los resultados del paso anterior.* Si en el paso anterior surgen malentendidos con definiciones, con alcance del tema, de condiciones o asuntos similares por parte de los expertos, deben ser resueltos sin lugar a dudas.
6. *Elicitación.* Este proceso tiene varias etapas en las cuales se realizan diferentes tareas y, si se tienen muchos parámetros a elicitar, debe dividirse de tal forma que el experto no se sienta abrumado por el trabajo. Se inicia siempre de una forma sencilla, registrando las impresiones del experto acerca del tema, determinando el nivel de claridad en el parámetro que se está estimando (no es raro ver sujetos que confundan la media con la mediana, aun en un proceso de capacitación). La contextualización es clave y es probable que se requiera material de apoyo a la mano.
7. *Análisis, agregación y documentación.* La documentación del proceso de elicitación es un paso obligatorio por parte del equipo coordinador del proceso. Todo el material utilizado y generado en este debe clasificarse adecuadamente junto con sus análisis. Los documentos utilizados tanto en la etapa de reclutamiento de expertos como entrenamiento de estos, las reuniones del equipo coordinador, las propuestas de análisis de resultados y la elicitación en sí misma (con grabaciones y material de video) deben hacer parte del material final del proceso. La parte de agregación de información de varios expertos es compleja, ya que depende del conocimiento que de los expertos tenga el grupo coordinador para determinar la forma de resumir las informaciones individuales (Genest y Zidek, 1986).

Las técnicas de elicitación presentadas aquí son útiles y permiten obtener información sobre un problema de una manera que puede pensarse que es objetiva. Esto

puede generar una confianza excesiva entre los investigadores, olvidando lo delicado de este proceso. Recomendamos, por lo tanto, tener claro lo siguiente:

- ¿Podemos dar garantía del nivel de experticia del experto?
- Tener siempre en mente al experto. ¿El experto entiende lo que le estamos preguntando? ¿Entiende claramente el problema?
- ¿Estamos manejando adecuadamente los tiempos para el proceso?
- Debemos validar cualquier procedimiento.

Garthwaite, Kadane, y O'Hagan (2004) y Jenkinson (2005) señalan varios aspectos que deben tenerse en cuenta en un proceso de elicitación:

- El objetivo es elicitación una distribución que represente el conocimiento presente del experto, y es útil tener un resumen sobre cuál es la justificación de este conocimiento.
- Cualquier interés personal o financiero que el experto pueda tener en las inferencias o decisiones que dependerán (aun marginalmente) de la distribución de este, debe ser declarado.
- Debe proporcionarse un entrenamiento para familiarizar al experto con las interpretaciones probabilísticas y con los conceptos y propiedades que se requieran en la elicitación. Es útil correr un ejercicio de elicitación ficticia para proporcionar práctica en el protocolo que el facilitador se propone utilizar.
- Debe llevarse un registro de la elicitación. Este debe contener, idealmente, todas las preguntas que fueron realizadas por el facilitador junto con las respuestas del experto, así como el proceso con el que se ajustó la distribución de probabilidad con estas respuestas.

Investigaciones han demostrado que, en muchas circunstancias, la gente espera que una muestra de una población represente todas las características esenciales de la población, aun si la muestra es pequeña. Esto es conocido como la ley de los pequeños números (Tversky y Kahneman, 1971; Sloman et al., 2003). El conocimiento de lo que ha ocurrido tiende a distorsionar la memoria y la gente se inclina a exagerar su probabilidad *a priori* para un evento que ha ocurrido. Los resultados muestran que los sujetos generalmente sobreestiman sus probabilidades *a priori* para eventos que ellos creen que han ocurrido, y subestiman sus probabilidades *a priori* para eventos que ellos creen que no han ocurrido.

Para poder usar estadística bayesiana, puede no ser necesario que una persona determine una distribución a priori completa y detallada, sino que debe ser capaz de

especificar alguna información de la cual una distribución a priori pueda ser construida. Sin embargo, especificar incluso una cantidad limitada de información acerca de una a priori puede ser difícil para mucha gente (Corless, 1972).

Oakley y O'Hagan (2007) consideran el caso de elicitación de la información a priori de un sujeto que solo puede dar información parcial. De todas formas, en un proceso de elicitación hay incertidumbre con relación a la información proporcionada por el experto. Ellos se limitan al caso en el cual el experto proporciona información sobre un θ observable. Chesley (1978) hace una comparación entre los métodos de elicitación. La técnica usada para la elicitación afecta cómo el sujeto mira el problema, la exactitud de sus respuestas y la consistencia de estas.

Corless realizó un estudio para responder las siguientes preguntas:

- ¿Están los auditores dispuestos a especificar información con la cual podamos construir las distribuciones a priori?
- ¿Asumiendo que los auditores especifiquen esta información, a qué se parecen las distribuciones a priori resultantes?
- ¿Qué tan bien reflejan estas distribuciones a priori las creencias del auditor acerca de las poblaciones a auditar?
- ¿Cómo son afectadas estas distribuciones a priori por las diferencias en la fortaleza del control interno y por las diferencias en los factores como la experiencia de los auditores?
- Si se usan estas distribuciones, ¿cómo se compara la evaluación bayesiana de la evidencia de auditoría con las evaluaciones intuitivas por parte del auditor de esa misma evidencia?

El experimento conducido por él consistió en presentarles dos casos, llamados A y B, a 88 contadores certificados. El caso A consistió en la descripción de una institución con un fuerte control interno y otra con un débil control. Para el experimento al conjunto de auditores se les describió el caso A y luego se les entregó el cuestionario. El primer conjunto de preguntas permitía establecer una distribución a priori y con el segundo se revisaban las probabilidades de esta. Lo mismo se hizo con el caso B.

Cuestionario utilizado para obtener información de las distribuciones a priori

1. Suponga que usted estuviera haciendo auditoría a varios clientes similares al hospital descrito en el caso A, ¿cuántos tendrían el mismo conjunto de circunstancias como este caso? Es claro que cada cliente puede tener diferentes tasas de error en la nómina, pero usted, probablemente, tiene alguna idea de cuáles pueden ser estas tasas de error.

- a) Especifique una mediana de la tasa de error tal que usted esperaría que la mitad de estos clientes tiene tasas de error mayores que esta mediana de la tasa de error y la otra mitad tendría tasas de error menores que esta mediana de la tasa de error. _____ %
- b) Especifique una tasa de error tal que usted esperaría que tres cuartas partes de estos clientes tengan tasas de error reales mayores que esta tasa de error estimada y que un cuarto de ellos tendrán tasas de error reales menores que ella. _____ %
(Esta respuesta debe dar un valor menor que el obtenido en la parte a)
- c) Especifique una tasa de error tal que usted esperaría que una cuarta parte de estos clientes tenga tasas de error reales mayores que esta tasa de error estimada y que tres cuartos de ellos tendrán tasas de error reales menores que ella. _____ %
(Esta respuesta debe dar un valor mayor que el obtenido en la parte a)
2. Si usted estuviera auditando 100 hospitales similares al descrito en caso A, los cuales tienen las mismas circunstancias, ¿cuántos de ellos esperaría usted que tendrían tasas de error reales en la nómina si estuvieran en cada uno de los siguientes intervalos de tasas de error?

Tabla 2.1. *Estimación del total para diferentes intervalos de tasas de error*

Parámetro	Valor
Menos del 1 %	
Entre el 1 % y el 2 %	
Entre el 2 % y el 5 %	
Entre el 5 % y el 10 %	
Más del 10 %	

(Sus respuestas deben sumar 100)

Las personas que trabajan en la elaboración de encuestas son conscientes del efecto que tienen tanto el cuestionario (preguntas) como la forma de realizarlo (persona o instrumento). Usualmente las personas que necesitan realizar las elicitaciones no son expertas en esta área y es necesario reconocer la importancia del tema.

Debemos distinguir dos procesos de elicitación, uno a partir de estudios previos y otro a partir de expertos. En muchos casos se trabaja con una combinación de expertos y estudios previos. El uso de estudios previos se ha refinado con la aparición del metaanálisis, el cual desarrolla métodos cuasestadísticos para combinar información de estudios de diversas variedades y procedencias, pero es aceptado entre miembros de algunas comunidades científicas. El uso de las dos fuentes de información, estudios previos y expertos, genera problemas que pueden ser difíciles de manejar, ya que se debe establecer la forma de ponderarlos.

2.2.5. Sobre los expertos

¿Quién es un experto? La caracterización de un experto es un tema difícil y ha sido motivo de muchas discusiones entre los investigadores (Kenney et al., 2011). Algunos han caracterizado a un experto como alguien con información relevante. También como alguien con información que está validada por un conocimiento profundo en el área específica de interés.

Adler y Ziglio en 1996 establecen cuatro criterios para decidir si alguien es un experto útil para un proceso de elicitación:

1. *Conocimiento y experiencia con los asuntos de interés.* Para determinar esto se puede recurrir inicialmente a la hoja de vida de un experto potencial, al reconocimiento por parte de sus colegas y a su autovaloración sobre su nivel de experticia en el tema.
2. *Capacidad y deseo de participar.* Muchas de las personas consideradas expertas manejan agendas ocupadas o viven en lugares distantes (aunque esto último se puede resolver con el uso de internet) o su ego hace que no consideren que deban compartir su conocimiento con personas que no son de su nivel. Con el ofrecimiento de honorarios que se ajusten a la posición de la persona, se puede ayudar a motivar al experto para que participe.
3. *Tiempo suficiente para participar.* Este punto está altamente ligado con el anterior. Los investigadores deben tener un buen estimativo de las horas de trabajo requeridas por el experto para entregar la información que se necesita y estar dispuestos a pagarlas, en especial con expertos de áreas profesionales como la salud, las ingenierías, el sector legal, etc.
4. *Habilidad de comunicación efectiva.* En procesos en los que es necesario interactuar intensamente con los investigadores sobre asuntos técnicos y que pueden generar dificultades de comprensión, es necesario que el experto sea capaz de comunicar sus creencias y justificarlas adecuadamente por escrito. En muchos casos, los investigadores deben establecer protocolos debidamente validados, de tal forma que en simulacros previos se determinen y resuelvan las dificultades que puedan surgir en el proceso.

2.2.6. Proceso de elicitación

DeWispelare et al. (1995) presentan 11 pasos en los cuales ellos dividen el proceso de elicitación a partir de expertos:

1. Determine los objetivos y propósitos del estudio.
2. Recrute a los expertos.

3. Identifique los temas e información que se necesitan. En el reporte de la DSTL (2014) hablan de la identificación y documentación de las variables para la elicitación.
4. Proporcione datos iniciales a los expertos.
5. Conduzca una sesión de entrenamiento en la elicitación.
6. Discuta y refine todos los asuntos involucrados.
7. Conduzca la elicitación. En el DSTL (2014) dividen esta etapa en cuatro más:
 - a) Elicite los juicios.
 - b) Ajuste una distribución a los juicios.
 - c) Presente la distribución ajustada al experto.
 - d) Refínela si es necesario.
8. Proporcione retroalimentación a los expertos.
9. Combine los juicios de los expertos (si se necesita).
10. Documente el proceso.

O'Connor y Remus (2005) realizaron un experimento donde hallaron que la técnica Feedback Intervention Theory (Teoría de Retroalimentación) mejora los juicios de los expertos.

Wallsten y Budescu (1983) presentan las condiciones para que un proceso de elicitación produzca una distribución apropiada. En teoría de la medición se mencionan con frecuencia las dificultades que tienen los procesos de cuantificación de sentimientos con relación a eventos externos y la determinación del error. El proceso de codificación debe garantizar condiciones básicas como confiabilidad y validez. La confiabilidad se mide como la correlación al cuadrado entre los valores observados de las probabilidades y las verdaderas probabilidades, $\rho_{p\pi}^2$. La probabilidad subjetiva es una variable aleatoria, p , que puede ser descompuesta como la verdadera probabilidad fija π y un error, e :

$$p = \pi + e \tag{2.3}$$

Los siguientes supuestos son estándares para este modelo:

1. $E(e) = 0$.
2. $\rho_{\pi e} = 0$.

3. Para cualquier par de mediciones independientes, los errores son incorrelacionados: $\rho_{e_i e_j} = 0$ para $i \neq j$.

4. $\rho_{\pi_i e_j} = 0$ para $i \neq j$.

Sea σ_e^2 la varianza del error. La raíz cuadrada de esta cantidad se conoce como el error estándar de la medición. Del modelo y los supuestos anteriores se tiene:

$$\sigma_p^2 = \sigma_\pi^2 + \sigma_e^2 \quad (2.4)$$

Así, el coeficiente de confiabilidad será:

$$0 \leq \rho_{\pi p}^2 = \frac{\sigma_\pi^2}{\sigma_p^2} = 1 - \frac{\sigma_e^2}{\sigma_p^2} \leq 1 \quad (2.5)$$

La confiabilidad está inversamente relacionada con el error de medición y es perfecta cuando $\sigma_e^2 = 0$. Ya que π no puede ser observado directamente, no podemos determinar $\sigma_{p_i}^2$. Esto puede resolverse parcialmente con métodos indirectos, por ejemplo, usando varios métodos de cuantificación.

La validez se define como la correlación entre dos procedimientos de cuantificación independientes, digamos ρ_{xy} .

Ayyub (2001) presenta una clasificación de la ignorancia que es importante considerar cuando se determina la claridad de un experto. La ignorancia puede ser consciente o ciega. La ignorancia ciega incluye conocimiento irrelevante que puede estar conformado por un conocimiento relevante y que es descartado o no considerado intencionalmente y por un conocimiento no confiable (prejuicios) o que no aplica al problema de interés.

Un elicitador subjetivo está bien calibrado si para cada probabilidad p , en la clase de todos los eventos en los cuales asigna una probabilidad subjetiva, la frecuencia relativa de ocurrencia es igual a p .

A pesar de que el concepto anterior es atractivo, en la práctica puede ser difícil de verificar, excepto en ciertas circunstancias en las que el elicitador permanentemente construye distribuciones de probabilidad subjetivas, como es el caso de los meteorólogos. En estos casos se pueden construir pruebas de tipo estadístico para medir el nivel de concordancia de lo elicitado con lo observado. En el nivel experimental se pueden establecer ambientes controlados en los cuales los elicitadores pueden ser evaluados; sin embargo, no hay garantía de que el resultado de un laboratorio pueda extrapolarse a un ambiente real.

Un proceso de elicitación de información mediante expertos se puede dividir en varios pasos. Los siete pasos para desarrollar un proceso formal de elicitación de Knol et al. (2010), información de expertos, son:

1. Caracterización de las incertidumbres

Las incertidumbres pueden ser cuantitativas o cualitativas, reducibles o permanentes, dependientes de diferentes métodos de medición o de diferentes valores personales.

Es necesario modelar la estructura de la incertidumbre. A los expertos se les puede pedir que construyan un modelo conceptual, que juzguen la verosimilitud de varios modelos propuestos o que proporcionen información sobre su conocimiento en un tópico particular que pueda usarse posteriormente para construir un modelo conceptual.

La naturaleza de la incertidumbre caracterizada puede deberse a un conocimiento parcial o incompleto, que se conoce como incertidumbre epistémica, o por las propiedades intrínsecas del sistema mismo, incertidumbre ontológica.

2. Alcance y formato de la elicitación

Los marcos de referencia temporal y espacial deben estar claramente establecidos.

3. Selección de los expertos

La búsqueda de expertos es un trabajo arduo ya que requiere tener una base amplia de hojas de vida de candidatos potenciales. Un experto no está obligado a participar en el estudio y debemos tener mucha claridad sobre los procesos de motivación para que lo haga, por ejemplo, considerando el pago de honorarios, la coautoría en publicaciones, etc.

4. Diseño del protocolo de elicitación

El número de entrevistas, las fechas y horarios, y el material a ser presentado al experto deben prepararse de antemano.

5. Preparación de la sesión de elicitación

La forma como se va a realizar cada sesión, si es presencial, telefónica o vía internet, debe estar claramente planificada. También, si se va a realizar una elicitación vía entrevista física o cuestionario, etc.

6. Elicitación de los juicios de los expertos

Ejecución de la elicitación planificada.

7. Posible agregación y reportes

Análisis de los resultados obtenidos, incluyendo la agregación de los resultados de los diferentes expertos. También viene un análisis de segmentación para determinar si hay heterogeneidad entre los expertos.

Christov et al. (2017) comparan varios métodos para llevar a cabo las elicitaciones y consideran:

- Entrevistas largas no estructuradas.
- Observaciones directas.
- Entrevistas semiestructuradas.
- Entrevistas semiestructuradas basadas en escenarios parciales.
- Entrevistas semiestructuradas basadas en escenarios completos.

El problema de la comparación de métodos de la forma planteada consiste en que es de carácter informal y no de tipo experimental. Aunque se puede entrever cuál método podría ser el mejor, al no controlar tantas fuentes de error que son inherentes a estos procesos, no hay ninguna garantía de la validez de los resultados. Bottomley y Doyle (2001) también han realizado comparaciones entre métodos.

2.2.7. Fallas en el proceso de elicitación

Los procesos de elicitación pueden tener errores de muy diversos orígenes:

Inexperiencia de los encargados del proceso

Los procesos de elicitación son muchas veces subvalorados y trivializados por parte de profesionales en áreas aun tan cercanas como la estadística. Usualmente, los profesionales de áreas donde se hagan estudios demoscópicos son conscientes de todas las dificultades técnicas de este proceso, pues en el levantamiento de información mediante encuestas se enfrentan problemas similares en algunos aspectos.

Mala elección de los expertos

Muchas de las personas que aparentemente son consideradas expertas en su área porque tienen una gran visibilidad profesional, por ejemplo políticos, pueden solo ser realmente expertos en el problema que se está tratando de resolver; por ejemplo, un parámetro muy específico o que la persona maneje más bien preconcepciones que le impiden ver el problema desde otras perspectivas. En muchas áreas sensibles podemos tener personas reconocidas pero no necesariamente que tengan las habilidades necesarias para participar en un proceso de elicitación. Muchos de los considerados expertos hablan desde sus intereses, a veces muy sesgados.

Mala preparación del proceso

Preguntas

- Debemos tener preguntas que apunten al objetivo central de la elicitación: Construir la distribución de credibilidad del experto.
- Toda pregunta debe apuntar al menos a un objetivo específico y debemos tener claridad sobre lo que se va a hacer con la respuesta.
- Hay que estar seguros de que el experto entiende todos los términos de la pregunta y no se va a sentir intimidado por esta.
- Debemos tener un protocolo claro, desarrollado de antemano, para procesar las respuestas.
- Un estándar para una buena pregunta es que el experto o expertos que la responden la deben entender de una manera consistente y que sea coherente con lo que el investigador espera que signifique.
- Debemos desarrollar preguntas que los expertos estén dispuestos a responder.
- Debemos asegurar que el experto conoce la respuesta.
- Debemos consultar con expertos en otras áreas la calidad de las preguntas que planeamos hacer, por ejemplo, un sicolingüista, un experto en entrevistas, etc.
- Hay algunas estrategias para chequear la validez de una pregunta: medir la consistencia de las respuestas del experto a la misma pregunta en dos puntos del proceso de elicitación.

Problema: determine el consumo diario promedio de gaseosas por parte de los estudiantes de pregrado de la universidad.

Posible pregunta: *¿Cuántos mililitros de gaseosa consume, en promedio, un estudiante en la universidad?* Esta pregunta posee una condición que puede generar confusión entre los expertos y es el manejo de las unidades propuestas. Un experto pudiera responder en términos de botellas o latas de gaseosas y posteriormente se hacen las conversiones necesarias.

- Evite preguntas que impongan condiciones indeseadas.
- Una pregunta a la vez.
- Sea cuidadoso con las preguntas que generen situaciones hipotéticas.
- Siempre tenga lápiz y papel.

- La respuesta 'no sé' hay que manejarla con cuidado, ya que el experto puede estar sugiriendo que no ha pensado lo suficiente en el problema planteado y, quizá, sea mejor darle tiempo para que elabore una respuesta coherente o entregarle información adicional. En otras ocasiones, el experto puede ser que no tenga ningún conocimiento acerca del tema planteado en la pregunta.

Tiempo de ejecución

Un problema común es la mala determinación del tiempo real que se necesita en el proceso de elicitación con el experto. No es raro que se requieran varios días, no obligatoriamente consecutivos. El experto muchas veces necesita tiempo para evaluar y recabar hechos que conoce acerca del problema que esté analizando. Siempre hay que tener claro que no se está evaluando el conocimiento del experto, sino que está en un proceso delicado que no debe generarle una reacción negativa hacia este.

Entorno

El entorno en el cual se realiza la elicitación debe proporcionar toda la comodidad posible para el experto, más aún en estos días, cuando se utilizan medios computarizados para realizar la entrevista que pueden no ser del agrado de todo el mundo.

Experto

Muchas veces los expertos pueden menospreciar de entrada el trabajo que se está realizando con ellos y tienden a dar respuestas rápidas por salir del paso, por ejemplo, si se siente evaluado. Otro caso especial es cuando el experto tiene una actitud negativa con respecto a las técnicas cuantitativas, en especial cuando caen en la cuenta de las dificultades para expresar en forma numérica, clara y coherente su conocimiento.

Las barreras para obtener una respuesta pueden ser:

1. El experto no tiene la información requerida.
2. No es capaz de recordar.
3. Tiene dificultad con el manejo del tiempo en que ocurren los hechos.
4. Tiene sesgos personales.

Entrevistador

La persona que realiza el proceso de elicitación debe generar toda la confianza de los expertos, ser una persona preparada para realizar entrevistas, que tenga una excelente vocalización, etc. Si un entrevistador intimida de alguna forma al experto, lo más seguro es que no se pueda llevar a cabo un proceso exitoso. Suponga una situación en la que se elicite a una persona ilegal para determinar su ingreso promedio

diario y quien realiza la entrevista le recuerde a una figura paterna o una autoridad; en este caso, podemos fracasar con el proceso.

Mala ejecución del proceso

Una de las fallas más comunes es la de subvalorar la complejidad del proceso y se descuidan o no se tienen en cuenta todos los detalles relacionados con este. A veces, en la planeación no realizamos simulacros y no logramos prever los problemas con la entrevista y las preguntas.

Boledo et al. (2009) presentan el siguiente protocolo para la determinación formal de los juicios de los expertos:

1. Entrenar los expertos en la cuantificación coherente de probabilidades.
2. Identificar y minimizar los sesgos de los expertos.
3. Definir y documentar sin lugar a dudas el problema a ser resuelto.
4. Proporcionarle al experto toda la información relevante que esté disponible.
5. Obtener la opinión de cada experto utilizando la técnica más apropiada, que puede ser diferente para cada uno.
6. Chequear y documentar la racionalidad y coherencia de cada experto en sus determinaciones.
7. Realizar una verificación final de todo el proceso, repitiéndolo si es necesario.

Debemos hacer una simulación completa del proceso de elicitación para encontrar posibles fallas de este.

Capítulo 3

Métodos de elicitación usando expertos

La captura de la información que un experto posee es una de las tareas más difíciles. Vamos a observar algunas técnicas para la captura de esta información teniendo siempre en mente que es un proceso extremadamente delicado y frágil, que nos puede llevar a equivocaciones.

En el proceso de elicitación debemos considerar dos casos: uno, en el que el sujeto no tiene restricciones en la forma de la densidad que representará su conocimiento a priori; este caso lo podemos llamar no paramétrico. En muchas circunstancias es conveniente, bien sea por cuestiones computacionales o por facilidad de interpretación (Hahn, 2006), modelar este conocimiento a priori mediante modelos paramétricos, por ejemplo, seleccionar la distribución Beta en el caso de una proporción, que es técnicamente deseable porque corresponde a la familia conjugada. Un argumento a favor puede ser que el problema de elicitación se simplifica, ya que uno se concentra en unos pocos parámetros, pero esto no es necesariamente cierto debido a que se requiere elicitar los parámetros de esta distribución, los cuales, si no tienen una interpretación intuitiva, pueden requerir las mismas técnicas que para el caso no paramétrico; este es el segundo caso. Se puede discutir si los sentimientos del experto realmente corresponden a un modelo paramétrico.

Slottje et al. (2008) clasifican las técnicas de elicitación en:

1. *Métodos en los que se fijan los valores.* El experto responde a preguntas como: ¿Cuál es la probabilidad de que el valor de la variable sea inferior (mayor) a un valor fijo arbitrario?
2. *Métodos en los que se fija el valor de la probabilidad.* El experto es cuestionado sobre el nivel de la variable que garantice que la probabilidad de que sea menor o igual a ese valor, sea un valor fijo y arbitrario π .

3. *Métodos en los que se utilizan intervalos.* Se le pide al experto separar en partes, equiprobables o no, el soporte de la distribución, por ejemplo, comenzando con la mediana y siguiendo con los cuartiles, etc.
4. *Métodos basados en loterías de referencia.* Al experto se le pide escoger entre dos apuestas. Una, con referencia a una lotería con una probabilidad conocida de ganar determinada y otra, donde el valor real de la cantidad está por debajo de cierto límite. La idea es buscar puntos de *indiferencia*, o sea, puntos de corte que generan una situación en la que al experto se le puede dar la lotería o la partición del espacio parametral por él definida, y no asume una preferencia por ninguna de las dos.

3.1. Formas de elicitación

El método de elicitación hace referencia a la forma o procedimiento que vamos a utilizar para cuantificar la información del experto.

En el proceso de elicitación podemos distinguir dos procedimientos de resolver el problema:

Métodos directos. Son apropiados para elicitar las distribuciones a priori de parámetros de interés que son intuitivos, por ejemplo, medidas de localización o proporciones.

Métodos indirectos. Son útiles en la mayoría de situaciones. El problema es elicitar la distribución poblacional y, a partir de esta, calcular la distribución a priori del parámetro de interés.

Además, podemos distinguir dos situaciones que aparecen en el trabajo aplicado:

Elicitación no paramétrica. En este escenario, el proceso de elicitación produce distribuciones a priori que no pertenecen a familias con estructuras específicas y determinadas por parámetros, aunque se pueden imponer algunas condiciones de suavidad, por ejemplo, la de continuidad o diferenciabilidad.

Elicitación paramétrica. En esta circunstancia, el proceso de elicitación está restringido a encontrar los valores de los parámetros de un modelo que se asume que representa bien el conocimiento existente; por ejemplo, en una proporción en la distribución beta se usa como un modelo que puede representar el conocimiento previo en muchos casos aplicados.

3.2. Métodos gráficos

La visualización de la incertidumbre acerca de un parámetro es un proceso relativamente directo y simple. Tiene la gran ventaja de ser muy intuitivo para la gran mayoría de expertos. Sin embargo, no todos los gráficos tienen la misma capacidad de representar estas incertidumbres.

3.2.1. Gráficos de distribuciones de probabilidad

Los métodos gráficos pueden ser utilizados tanto en etapas exploratorias (una primera aproximación a la distribución a priori) como para representar la información definitiva. La gran ventaja que tienen es que el usuario puede expresar sus creencias con relativa facilidad y realizar correcciones de forma rápida, lo cual es una ventaja, en especial cuando queremos calibrar las obtenidas por otros métodos.

Los métodos gráficos se pueden dividir en dos áreas:

1. **Gráficos de densidades:** este es un método sencillo que puede realizarse a mano alzada o respondiendo a preguntas específicas sobre probabilidades de regiones del espacio parametral. Puede expresarse mediante histogramas o mediante gráficas a mano alzada continuas. Mazzuchi et al. (1991) y Van Noortwijk et al. (1992) presentan una aplicación del método gráfico para un problema de confiabilidad basado en el uso de histogramas. Johnson et al. (2010) utilizan un método gráfico para determinar la densidad a priori. Es fácil hacer representaciones bidimensionales.

Esta aproximación permite elaborar plantillas que involucren expresiones verbales sobre incertidumbre, lo cual es de gran ayuda al experto para expresar sus creencias. Una vez tenemos la gráfica, es posible normalizarla de tal forma que se garanticen las condiciones de función densidad de probabilidad. Otra gran ventaja es que podemos buscar una distribución paramétrica que se ajuste a esta densidad de tal forma que simplifique el análisis posterior.

Proponemos un procedimiento gráfico que puede ayudar a una persona a expresar su opinión por medio de una distribución de probabilidad que refleja su nivel de seguridad sobre los posibles valores que puede tomar.

Determinación de la distribución *a priori* a mano alzada

El uso de técnicas a mano alzada no ha sido muy popular entre los matemáticos y fue Tukey quien propuso el uso de gráficos como base de un primer análisis de los datos. Un argumento para esta aproximación se basa en la gran capacidad que tienen los humanos para detectar patrones en situaciones complejas. Uno

de los problemas que enfrentamos es que los “datos” están dentro de la cabeza del experto y no están en formas numéricas, sino con sus creencias y sentires, que pueden cambiar como lo hacen sus procesos mentales. Aquí se propone la técnica de elicitación de curvas de creencia para determinar el nivel de “fe” en que un parámetro, desconocido pero de interés, tome ciertos valores con mayor o menor seguridad por parte del experto.

Mediante la ayuda de una plantilla como la presentada en la figura 3.1 podemos pedirle al usuario que determine el nivel de seguridad que tiene acerca de un subconjunto de valores posibles del parámetro. Es mucho más sencillo para el usuario tener una escala ordinal estableciendo su nivel de seguridad, que determinar un posible valor asociado con algún subconjunto. Esta plantilla utiliza expresiones verbales sobre incertidumbre, las cuales no son de carácter universal. Existe poco trabajo que trata de la cuantificación de estas expresiones, como se muestra abajo.

Una ventaja de este procedimiento es que la persona puede realizar elicitación a mano alzada de una distribución poblacional o de un parámetro particular de interés. La primera se puede usar en situaciones en las cuales el parámetro de interés no sea intuitivo, por ejemplo, pensar sobre alguna medida como curtosis.

Estando la escala respuesta en forma nominal-ordinal, no es clara la posición de cada expresión en el eje vertical; este es el principal problema de esta metodología.

Al menos por dos razones los analistas requieren que los juicios acerca de incertidumbre sean numéricos:

- a) Se cree que las expresiones numéricas son comunicaciones precisas e inequívocas que permiten el cálculo del valor esperado o la utilidad esperada, mientras que el lenguaje natural es vago, sujeto a diferentes interpretaciones por diferentes personas y no es útil para cálculos claros.
- b) La calidad de las expresiones numéricas pueden ser evaluadas mientras que la de las lingüísticas no.

Las expresiones verbales sobre incertidumbre pueden ayudar a validar las expresiones numéricas de estas. La gente prefiere la comunicación verbal sobre la numérica cuando quiere expresar su incertidumbre, pero prefiere recibirla en forma numérica. Una de las motivaciones más fuertes para usar representaciones de incertidumbre basadas en el lenguaje es evitar el problema de la “falsa precisión” asociado con las probabilidades establecidas numéricamente. Mosteller y Youtz (1990) desarrollaron un estudio asociado con expresiones verbales en inglés, pero, a pesar de que las palabras puedan traducirse de una manera relativamente directa, esto no necesariamente implica que los valores asociados por ellos correspondan en español. Además, esta investigación consideró que cada persona entrevistada asociaba un único valor con cada expresión verbal considerada por ellos y un rango que determinaba el grado de certidumbre acerca de este valor. Esto puede no ser apropiado, ya que un sujeto asociado con una expresión verbal, por ejemplo, “probablemente lloverá mañana”, puede no identificar un único valor de probabilidad asociado con la palabra *probablemente*. Por lo tanto, proponemos considerar, en su lugar, una elicitación de una distribución de probabilidad del nivel de certidumbre asociada con esta palabra.

La comunicación de diferentes niveles de incertidumbre sobre la ocurrencia de eventos particulares se realiza utilizando expresiones comunes del lenguaje, tales como posiblemente, no creo que, estoy seguro, etc. En la comunicación científica y técnica también son usadas estas expresiones, dando origen a confusiones entre los diferentes actores, ya que la identificación de una expresión verbal con un único valor numérico (probabilidad) no es clara.

Para que sea útil la cuantificación de la distribución de probabilidad de los valores numéricos de una expresión verbal, no podemos considerar solo la proporcionada por un individuo. Para esto, requerimos considerar las distribuciones proporcionadas por un grupo de individuos, generando una distribución de probabilidad representativa.

Para esto realizamos el siguiente experimento: obtuvimos una muestra de sujetos que tuvieran un conocimiento de estadística en el nivel de los cursos dictados

en pregrado. Consideramos 30 individuos; a cada uno se le presentaron cinco expresiones verbales, se le explicó objetivo del experimento y el procedimiento que debían seguir para elicitar sus distribuciones. Las expresiones fueron presentadas a cada sujeto luego de una permutación aleatoria.

Una vez tenemos las 30 funciones de probabilidad halladas a mano alzada, debemos desarrollar un procedimiento para obtener una función que sea representativa. Una posible aproximación es la siguiente:

- a) Determine una rejilla del intervalo $[0, 1]$ (el valor numérico de una expresión verbal), digamos $\pi_1, \pi_2, \dots, \pi_M$.
- b) Para el sujeto i , determinamos $f_i(\pi_1), f_i(\pi_2), \dots, f_i(\pi_M)$.
- c) Para cada punto π_j , calculamos la mediana de $\{f_1(\pi_j), f_2(\pi_j), \dots, f_{30}(\pi_j)\}$, denotado por $\hat{f}(\pi_j)$.
- d) Mediante alguna técnica de suavización ajustamos una curva suave a los puntos $\{\hat{f}(\pi_1), \hat{f}(\pi_2), \dots, \hat{f}(\pi_M)\}$

Las frases que involucraban términos sobre incertidumbre fueron:

- a) “El meteorólogo afirmó en el noticiero que era probable que mañana lloviera”.
- b) “Es raro que un paciente desarrolle la enfermedad X si es mayor de Y años”.
- c) “A menudo un hijo de un fumador desarrolla el hábito de fumar”.
- d) “Es posible que haya vida extraterrestre”.

3.3. Algoritmo para la determinación de la a priori paramétrica

- a) Obtenga la distribución a mano alzada, digamos $\xi_A(\theta)$.
- b) Discretice $\xi_A(\theta)$ en una rejilla $\Theta_D = \{\theta_1, \theta_2, \dots, \theta_K\}$ y evalúe $\xi_A(\theta)$ en estos puntos.
- c) Muestree con reemplazo de Θ_D M valores (M grande) con probabilidades proporcionales a $\{\xi_A(\theta_1), \xi_A(\theta_2), \dots, \xi_A(\theta_K)\}$.
- d) Calcule los momentos de esta distribución M_1, M_2, \dots, M_k .
- e) Iguale los momentos hallados en el punto anterior a los k momentos de la distribución paramétrica, de modo que, al resolver estas ecuaciones, se encuentren de forma única los parámetros de la distribución objetivo.

Para medir la calidad de la aproximación, podemos usar:

$$\int_{\Theta} |\xi(\theta) - \xi_A(\theta)| \xi_A(\theta) d\theta \approx \frac{1}{K} \sum_{i=1}^K |\xi(\theta_i) - \xi_A(\theta_i)| \frac{\xi_A(\theta_i)}{\sum_{j=1}^K \xi_A(\theta_j)} \quad (3.1)$$

La siguiente tabla muestra algunos datos estadísticos correspondientes a 43 expresiones probabilísticas usadas por Mosteller y Youtz (1990).

Tabla 3.1. *Media, mediana, desviación típica y rango de la cuantificación correspondiente a 43 expresiones probabilísticas*

Frase	Número de respuestas	Media	Mediana	Desviación típica	Rango
Altamente probable	187	0.89	0.90	0.04	0.60-0.99
Muy verosímil	185	0.87	0.90	0.06	0.60-0.99
Muy probable	187	0.87	0.89	0.07	0.60-0.99
Bastante verosímil	188	0.79	0.80	0.10	0.30-0.99
Usualmente	187	0.77	0.75	0.13	0.15-0.99
Buena posibilidad	188	0.74	0.75	0.12	0.25-0.95
Predecible	146	0.74	0.75	0.20	0.25-0.95
Verosímil	188	0.72	0.75	0.11	0.25-0.99
Probable	188	0.71	0.75	0.17	0.01-0.99
Más bien verosímil	188	0.69	0.70	0.09	0.15-0.99
Una buena oportunidad	188	0.67	0.70	0.12	0.25-0.95
Es verosímil	188	0.66	0.70	0.12	0.15-0.95
Algo verosímil	187	0.59	0.60	0.18	0.20-0.92
Un poco más que casualidad	187	0.58	0.60	0.06	0.45-0.89
Quizás	124	0.58	0.60	0.11	0.10-0.80
Un poco más que la mitad de las veces	188	0.55	0.55	0.06	0.45-0.80
Una pequeña ventaja	187	0.55	0.55	0.08	0.05-0.75
Una ligera ventaja	188	0.51	0.50	0.13	0.20-0.85
Carisellazo	188	0.50	0.50	0.00	0.45-0.52
Posibilidades casi iguales	186	0.47	0.50	0.17	0.05-0.90
Un poco desventajoso	188	0.45	0.45	0.04	0.05-0.50
Las apuestas están ligeramente en contra	185	0.45	0.45	0.11	0.10-0.99
No muy parejo	180	0.44	0.45	0.07	0.05-0.60
Inconcluso	153	0.43	0.50	0.14	0.01-0.75
Incierto	173	0.40	0.50	0.14	0.08-0.90
Posible	178	0.37	0.49	0.23	0.01-0.99
Algo inverosímil	186	0.31	0.33	0.12	0.03-0.80
Apenas inverosímil	187	0.25	0.25	0.11	0.02-0.75
Más bien inverosímil	187	0.24	0.25	0.12	0.01-0.75
Posiblemente inverosímil	187	0.21	0.20	0.10	0.01-0.75
No muy probable	187	0.20	0.20	0.12	0.01-0.60
Inverosímil	188	0.18	0.16	0.10	0.01-0.45
No mucha posibilidad	186	0.16	0.15	0.09	0.01-0.45
Raro	188	0.16	0.15	0.08	0.01-0.47
Escasamente posible	180	0.13	0.05	0.17	0.01-0.60
Difícilmente posible	184	0.13	0.05	0.16	0.01-0.50
Improbable	187	0.12	0.10	0.09	0.01-0.40
Bastante inverosímil	187	0.11	0.10	0.08	0.01-0.50
Muy inverosímil	186	0.09	0.10	0.07	0.01-0.50
Muy raro	187	0.07	0.05	0.07	0.01-0.30
Altamente improbable	181	0.06	0.05	0.05	0.01-0.30

Witteman y Renooij (2003) presentan una escala verbal mucho más pequeña, construida a partir de expertos.

Tabla 3.2. *Escala verbal para siete expresiones probabilísticas*

Expresión	Probabilidad
Cierto (seguro)	100
Probable	85
Se espera	75
50-50	50
Incierto	25
Improbable	15
(casi) imposible	0

En trabajos efectuados sobre el tema, Theil (2002) discute sobre las variaciones en las cuantificaciones en términos probabilísticos de expresiones verbales. Por ejemplo, la palabra “posible” toma valores de 0.55, 0.27, 0.33, 0.37, 0.44 y 0.47 en diferentes trabajos. Ferson et al. (2015) realizan un estudio sobre expresiones que manifiestan incertidumbre y presentan una clasificación de cada expresión en uno de tres tipos: simétrica, asimétrica o rango.

La tabla 3.3 presenta los valores medios obtenidos en una investigación adelantada entre 140 sujetos para la cuantificación de expresiones de incertidumbre. Hay que recordar que estos trabajos se han realizado con frases o palabras en inglés y, por lo tanto, para las expresiones en español no se ha desarrollado ningún trabajo hasta el presente.

Tabla 3.3. *Valores medios estimados para distintas expresiones*

Palabra	Valor
Absolutamente imposible	0.015
Raramente	0.119
Muy improbable	0.159
Raro	0.196
No muy probable	0.214
Bastante improbable	0.271
Algo improbable	0.318
Incierto	0.404
Algo menos que la mitad de las veces	0.439
Un carisellazo	0.498
Algo más que la mitad de las veces	0.456
Más que parejo	0.611
Algo posible	0.677
Buen chance	0.723
Es posible	0.769
Muy probable	0.835
Altamente probable	0.880
Casi seguro	0.913
Absolutamente seguro	0.998

Moore (1977) presenta otras expresiones con valores diferentes, lo cual muestra lo difícil que es este problema. Fillenbaum et al. (1991) examinan cómo los significados de palabras que denotan incertidumbre en inglés, tales como *likely* (posiblemente) o *probable* (probable), son alterados por modificadores como *very* (muy). Además, contemplan la distribución de probabilidad del nivel de incertidumbre que tiene cada palabra de las consideradas.

Determinación de las distribuciones de probabilidad

2. **Gráficos de distribuciones acumuladas:** pueden acompañar métodos que hagan referencia a preguntas del estilo de: “¿Qué porcentaje de unidades tienen un valor en la variable de interés menor o igual que un valor x ?” La experiencia ha mostrado que este procedimiento es más difícil que el graficar las densidades, aunque es muy bueno para comparar visualmente la información proporcionada por diversos expertos o para ver si la distribución subjetiva puede aproximarse mediante un modelo paramétrico o los cambios que tiene un experto en distintas fases del proceso de elicitación. Es muy útil para representar la información obtenida mediante otros procedimientos.

¿Cómo llevar a una función de probabilidad nuestra gráfica a priori?

Como una primera aproximación al proceso de determinación de la distribución a priori, podemos hacer una representación gráfica de nuestras creencias. Los gráficos de densidades de información subjetiva a mano alzada caen en la categoría de determinación de creencias (Benson et al., 1995). Kloprogee et al. (2007) discuten algunas posibles desventajas en la expresión gráfica de incertidumbres. Entre estas se tiene que el lector de las gráficas puede interpretarlas a su manera y, de esta forma, distorsionar parte del mensaje transmitido en ellas. Obviamente, empezamos graficando en un eje horizontal un rango de valores del parámetro a determinar, de modo que estemos absolutamente seguros de que el valor cae ahí. Por ejemplo, si estamos elicitando la estatura de alguien que conocemos, pudiéramos estar absolutamente seguros de que esa persona tiene menos de 190 cm y más de 175 cm. Podríamos, por lo tanto, concentrarnos en este intervalo.

Este es un método muy intuitivo y puede utilizarse para elicitación de parámetros asociados con características únicas, tales como la edad de una persona determinada, la probabilidad de que una misión espacial falle en cierta etapa y de que funcione una solución completamente nueva e irrepetible a un problema ambiental.

Un problema que aparece con cierta frecuencia al utilizar este procedimiento es la tendencia de los expertos a graficar funciones unimodales. Estos son problemas que, a veces, se pueden resolver informándoles a los expertos la posibilidad de que en rangos del espacio parametral, se pueden tener indeterminaciones sustanciales entre valores muy cercanos y esto daría origen a distribuciones escalonadas.

Ilustración: cálculo de la edad de una persona

En nuestra sociedad es considerado de mala educación preguntar la edad de una persona. El día en que conoce a alguien, usted más o menos puede calcularle la edad. Este proceso se hace de una manera inconsciente y usualmente llega a un número que aproxima sus creencias sobre la posible edad. Para esto usa la información recolectada previamente sobre ella, por ejemplo, si esta persona tiene una apariencia determinada, si se viste de cierta forma, si se graduó del colegio en cierta época, etc. Si dos personas tienen que calcular la edad de este sujeto, puede suceder que no coincidan en sus resultados, pero no se puede decir cuál de las dos está equivocada (o si las dos lo están), solo hasta el momento en que se conozca la verdadera edad de la persona en cuestión.

Un retrato o una foto puede ser el punto de partida para elicitación de la edad de una persona, por lo que suponga que se tiene a la mano la foto de un individuo al que se le desea estimar su edad vía elicitación. Cuando vemos la foto, asociamos intuitivamente situaciones similares de las que tengamos conocimiento; en este caso, buscamos en nuestros recuerdos personas que hayamos conocido que se parezcan al sujeto de la

foto. Si tenemos a alguien que sea un buen referente, tendemos a asociar su edad con la del sujeto de interés, haciendo ajustes que consideramos adecuados. Ajustando el eje X en la plantilla que se presenta en la figura 3.1, se puede elicitar la distribución de probabilidad que nos refleja la incertidumbre que tenemos sobre la edad de una persona. Por ejemplo, en el caso de la estatura de una persona, podríamos escoger un intervalo que vaya de 0 metros a 3 metros. Alguien puede darnos un rango más fino (estrecho), dado su conocimiento a priori. Continuando con la elicitación de la edad de una persona que estamos seguros de que se encuentra entre los 35 y los 56 años, podemos construir también dicha plantilla a mano alzada y su apariencia puede ser como se muestra en la siguiente figura.

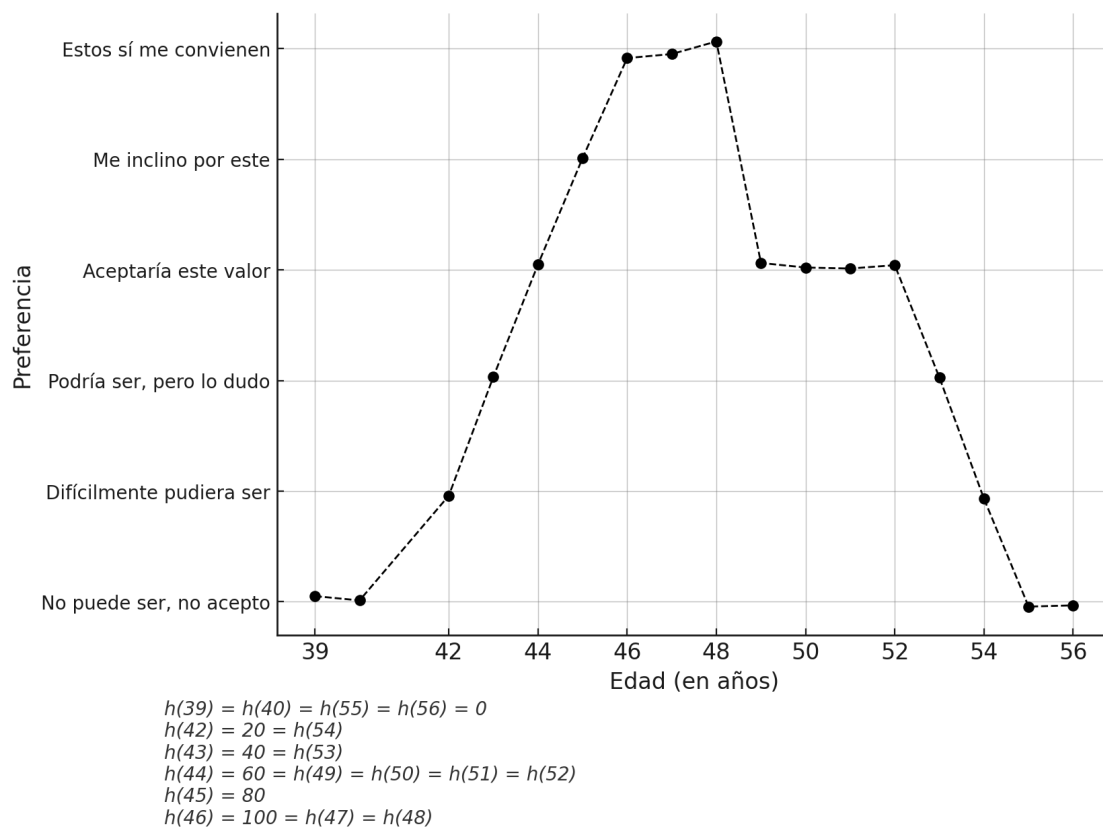


Figura 3.2. *Distribución a priori a mano alzada para la edad de una persona*

Observe que el proceso se puede realizar a mano alzada en sus etapas iniciales, ya que este exige revisar permanentemente los juicios y las opiniones del experto. La altura de la densidad en un punto es proporcional a la verdadera altura.

El cálculo de las características típicas de una distribución, tales como el promedio, la varianza y los percentiles, etc., se puede realizar con diferentes aproximaciones. El más simple es considerar la edad discreta tradicional, o sea, solo tenemos edades enteras y en este caso podemos usar R para hallar la media y la varianza, así:

```

# Ejemplo sobre la edad de elicitada a mano

edad<-39:56
h.edad<-c(0,0,1.5,3,6,9,12,15,15,15,9,9,9,9,6,3,0,0)

prob.edad<-h.edad/sum(h.edad)

(media.elicitada<-sum(edad*prob.edad))
[1] 47.69136
(var.elicitada<-sum((edad-media.elicitada)^2*prob.edad))
[1] 9.670172

# Solución vía simulación
muestra<-sample(edad,10000,replace=T,prob=prob.edad)
mean(muestra)
[1] 47.7245
var(muestra)
[1] 9.562356
median(muestra)
[1] 47
quantile(muestra, probs=c(0.05,0.5,0.95))
5% 50% 95%
43 47 53

plot(edad,cumsum(prob.edad),type='s',ylab='Probabilidad',
      xlab='Edad')
points(edad,pnorm(edad,mean=media.elicitada,sd=sqrt(var.elici
tada)),type='l')
title(main='Función de distribución acumulada')

```

Función de distribución acumulada

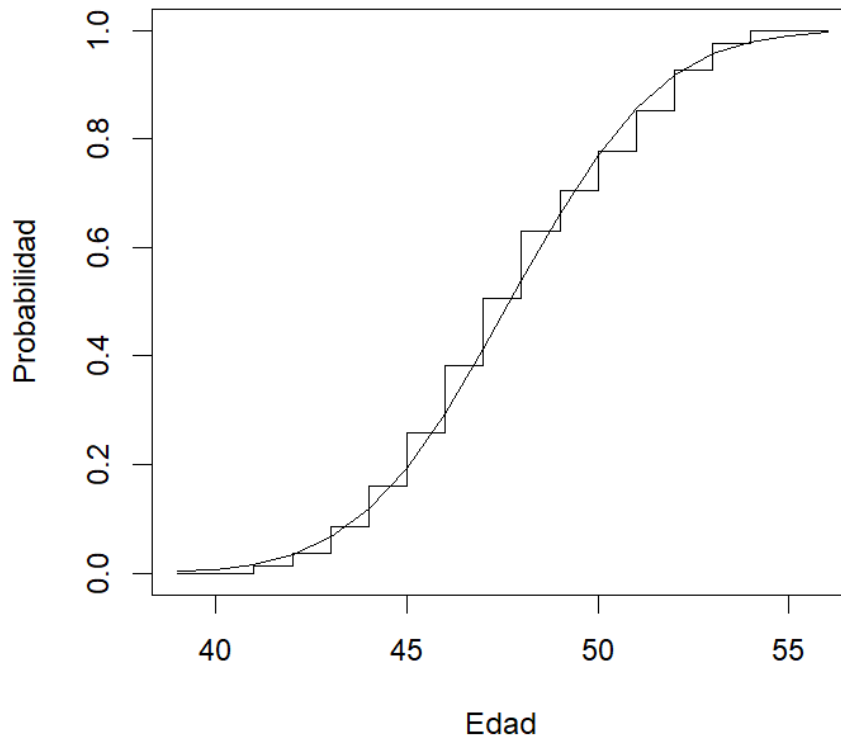


Figura 3.3. *Función de distribución de la edad de la persona que fue elicitada a mano alzada junto con la función de distribución de una normal con media 47.7245 y varianza 9.670172*

3.4. Métodos directos

Una forma de asignación de probabilidades es la interrogación directa acerca del parámetro de interés o representación gráfica.

González-Vallejo et al. (1994) afirman que la forma de determinar probabilidades depende de si estas se expresan verbal o numéricamente. La búsqueda de expresiones verbales que muestren algún nivel de incertidumbre es una tarea difícil y es dependiente del idioma en que se esté realizando la evaluación con una fuerte carga regional; por ejemplo, el idioma español presenta grandes diferencias entre los colombianos y los argentinos. Para complicar el problema, existen diferencias regionales que habría que considerar. Druzdzel (1989) presenta un inventario de más de 150 frases en inglés que reflejan algún grado de incertidumbre.

3.5. Obtención directa de probabilidades

El cálculo directo de probabilidades sobre eventos ha sido una de las alternativas más utilizadas en la literatura. Ludke et al. (1977) comentan que algunos investigadores han encontrado que las probabilidades halladas mediante este método tienden a sobreestimarse si ellas son pequeñas y a subestimarse si son grandes. Se han propuesto algoritmos que garantizan las condiciones teóricas y se discuten cuestiones de consistencia en los resultados que se obtienen.

3.5.1. Histogramas

El método del histograma es uno de los más intuitivos para que un sujeto determine su distribución a priori.

3.5.2. Método de intervalo fijo

El rango posible (este debe ser determinado por el experto) se divide en un número de intervalos de igual amplitud. Estos intervalos se le presentan al asesor, quien determina cuál es la probabilidad correspondiente a cada uno. Usualmente, el ejercicio comienza mediante la ordenación de los intervalos, del más probable al menos probable, y luego se determinan los valores de probabilidad.

3.5.3. Método de intervalo variable con particiones equiprobables o método de bisección

En este método se le pide al experto que determine intervalos que correspondan a probabilidades predeterminadas. Una forma posible es preguntarle un valor θ_0 , tal que en su opinión el verdadero valor del parámetro que esté por encima sea igual a la probabilidad que está por debajo (percentil 50). Este procedimiento funciona igual si consideremos una variable poblacional tal como la estatura o los ingresos. Al experto se le pide que imagine que el verdadero valor de θ está por debajo de θ_0 . Nos debe presentar, entonces, un valor θ_1 , tal que él crea que es igualmente posible que el verdadero valor de θ esté por debajo de este valor y es igual a la probabilidad que esté entre θ_1 y θ_0 . Este proceso se repite para refinar la distribución tanto como se desee. Obviamente, se encuentran puntos en los cuales el experto es incapaz de partir un intervalo en intervalos equiprobables. Este método también es conocido como *método de bisección sucesiva*; tiene la ventaja de estar todo el tiempo creando pares de intervalos equiprobables. Un problema que puede aparecer es que si hay una equivocación en una de las particiones, el resto del proceso que involucre esta partición particular se verá afectado.

Garthwaite et al. (2005) utilizan los terciles, los cuales corresponden a los percentiles $33\frac{1}{3}$ y $66\frac{2}{3}$, con el argumento de que estudios empíricos muestran que tienen mejores propiedades.

Un problema con este procedimiento es que, al iniciar con la mediana los expertos pueden terminar con soportes que sean más estrechos que los que ellos hubieran seleccionado con otros procedimientos.

Particiones no equiprobables han sido también propuestas (Firmino y López Droguett, 2015). Estas pueden ser atractivas en proceso de elicitación, ya que el experto puede desprenderse de las apuestas en forma explícita. Puede tener dificultades en la selección de los puntos de las particiones y en la determinación de qué tanto más verosímil es una región con respecto a la otra.

Ejemplo del método de bisección. Un ejemplo de esta aproximación es presentada en Lapin (1994):

El presidente de una compañía de alimentos procesados desea obtener la distribución de probabilidad para la demanda de un nuevo *snack*. Esta será usada para determinar si se lanza o no el nuevo producto en el mercado. Un analista estadístico le hace una serie de preguntas.

Pregunta: ¿Qué piensa usted del mayor y del menor valores posibles para la demanda?

Respuesta: con seguridad la demanda superará las 500.000 unidades, pero yo establecería un techo de 3.000.000 de unidades. No creo que aun bajo las condiciones más favorables se venda más de esta cantidad.

P. Está bien, hemos ya determinado un rango para la posible demanda. Ahora, quisiera que me dijera, ¿cuál valor de la demanda divide la demanda en dos rangos igualmente posibles? Por ejemplo, ¿creería usted que la demanda estará por debajo de 2.000.000 de unidades o por encima de este valor con igual posibilidad?

R. No. Yo escogería mejor 1.500.000 como un punto de 50-50 chances de estar por encima o por debajo.

P. Muy bien. Ahora consideremos que la demanda estuviera por debajo de 1.500.000. Si la demanda estuviera en algún punto entre 500.000 y 1.500.000, ¿apostarías que está por debajo o por encima de 1.000.000?

R. Por encima. Yo diría que una demanda de 1.250.000 sería un punto de división realista.

P. Utilizaremos ese punto como un 50-50 chances. Ahora, hagamos lo mismo para una demanda superior a 1.500.000.

R. Si yo fuera a escoger un punto, escogería 2.000.000. Yo creería que la demanda es igualmente probable que caiga entre 1.250.000 y 2.000.000, que por encima de este rango.

P. Excelente. Estamos progresando. ¿Para refinar más estos valores, pensemos si es igualmente posible que la demanda caiga entre 1.250.000 y 2.000.000 que por fuera de este rango?

R. No. Yo creo que es más posible que caiga en el interior de este intervalo que por fuera. Creo que estoy siendo inconsistente.

P. Sí, lo está siendo. Vamos a remediarlo. ¿Piensa que debemos elevar o disminuir el valor 1.250.000?

R. Bajarlo de 2.000.000 a 1.900.000.

P. Chequeemos si esto perturba las otras respuestas. ¿Piensa que el valor 1.500.000 divide la demanda de 1.250.000 a 1.900.000 en dos partes igualmente posibles?

R. Sí, estoy satisfecho con ese valor.

P. Solo unas pocas preguntas más. Suponga que la demanda estuviera por encima de 1.900.000. ¿Cuál valor dividiría el intervalo en dos intervalos igualmente posibles?

R. Yo diría que 2.200.000.

P. Bien. Ahora, si la demanda estuviera entre 2.200.000 y 3.000.000, ¿dónde estaría el punto 50-50?

R. Yo creo que 2.450.000 unidades sería el punto 50-50.

P. ¿Qué hay si la demanda se asumiera por debajo de 1.250.000?

R. Ensayemos con 1.100.000 unidades.

P. ¿Y cuando la demanda está entre 500.000 y 1.100.000?

R. Yo creería que es muy posible que esté cerca del límite superior. Yo apostaría a que es 950.000 unidades.

Tabla 3.4. *Cuantiles estimados vía método de bisección*

Cuantil	Cantidad
0	500.000
0.0625	950.000
0.125	1.100.000
0.25	1.250.000
0.50	1.500.000
0.75	1.900.000
0.875	2.200.000
0.9375	2.450.000
1.000	3.000.000

En R podemos utilizar el siguiente código para realizar las gráficas de la función de distribución acumulada y la densidad.

```
temp<-scan()
0      500000
0.0625 950000
0.125  1100000
0.25   1250000
0.50   1500000
0.75   1900000
0.875  2200000
0.9375 2450000
1.000  3000000

temp<-matrix(temp,ncol=2,byrow=T)

plot(temp[,2],temp[,1],type='b',ylab='Cuantil',xlab='Cantidad')
abline(h=temp[,1],lty=2)

densidad<-function(temp,marcaX='XXX',marcaY='YYY'){
k<-nrow(temp)-1
altura<-NULL
for(i in 1:k){
  altura<-c(altura,(temp[i+1,1]-temp[i,1])/(temp[i+1,2]-temp[i,2]))
}
plot(0,0,ylim=c(0,max(altura)),xlim=c(temp[1,2],temp[k+1,2]),
ylab=marcaY,xlab=marcaX)

for(i in 1:k) polygon(c(temp[i,2],temp[i,2],temp[i+1,2],
temp[i+1,2],temp[i,2]),c(0,altura[i],altura[i],0,0),col='grey')
}

densidad(temp,marcaX='Cantidad',marcaY='Densidad')
```

3.5.4. Función de distribución acumulada

La construcción se realiza respondiendo preguntas del tipo: “¿Qué porcentaje de valores del parámetro de interés está por debajo del valor dado, digamos θ_i ?”. Esto se hace para un conjunto de valores $\theta_1, \theta_2, \dots, \theta_k$. A partir de esto se ajusta un modelo paramétrico o un modelo empírico.

En el caso de la elicitación vía cuantiles, Lau et al. (1999) consideran la construcción de la distribución subjetiva completa usando la curva de Tocher, pero reconocen que esta aproximación puede presentar problemas serios, ya que sus resultados son más exactos. Sin embargo, ellos presentan el cálculo de los momentos a partir de interpolaciones lineales sobre la base del resultado.

$$\mu'_k = E(X^k) = \int_0^1 [F^{-1}(p)]^k dp \quad (3.2)$$

Los momentos centrales pueden construirse a partir de los momentos alrededor de cero, como:

$$\begin{aligned} \mu_2 &= \mu'_2 - (\mu'_1)^2 \\ \mu_3 &= \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3 \\ \mu_4 &= \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4 \end{aligned} \quad (3.3)$$

Asumamos que el proceso de elicitación produjo los cuantiles y las probabilidades (p_i, C_i) para $i = 1, \dots, L$. La línea que corresponde al i -ésimo segmento para la función $\hat{F}^{-1}(p)$ es:

$$\begin{aligned} \hat{F}^{-1}(p) = C &= C_{i-1} + \frac{C_i - C_{i-1}}{p_i - p_{i-1}} \times (p_i - p_{i-1}) \\ &= C_{i-1} + S_i \times (p_i - p_{i-1}) \end{aligned} \quad (3.4)$$

donde $S_i = \frac{C_i - C_{i-1}}{p_i - p_{i-1}}$ es la pendiente del i -ésimo segmento de $\hat{F}^{-1}(p)$. Por tanto,

$$\int_{p_{i-1}}^{p_i} [\hat{F}^{-1}(p)]^k dp = \int_{p_{i-1}}^{p_i} [C_{i-1} + S_i \times (p - p_{i-1})]^k dp = \frac{C_i^{k+1} - C_{i-1}^{k+1}}{S_i(k+1)} \quad (3.5)$$

Así

$$\begin{aligned} \mu'_k &\approx \sum_{i=1}^{L+1} \int_{p_{i-1}}^{p_i} [\hat{F}^{-1}(p)]^k dp \\ &\approx \sum_{i=1}^{L+1} \frac{C_{(i)}^{k+1} - C_{(i-1)}^{k+1}}{S_i(k+1)} \end{aligned} \quad (3.6)$$

3.5.5. Función densidad de probabilidad

Este procedimiento es más complejo que el anterior, ya que determinar valores de densidad puede no ser fácil y algunas personas prefieren trabajar con histogramas.

- No paramétrica. En este caso, se utiliza el modelo empírico (el elicitado realmente). Forester et al. (2004) presentan el caso de la elicitación de la probabilidad de falla mediante preguntas sobre la peor probabilidad posible y se asocia con el percentil 99 %, el mejor caso (la probabilidad de riesgo más baja) que se asocia y la mediana. Además, se determinan los valores correspondientes a los cuartiles y a los percentiles 10 % y 90 %.
- Paramétrica. Este procedimiento exige la determinación previa de un modelo paramétrico, por ejemplo, una distribución beta para el caso de elicitar una proporción, que sea apropiada para representar la información previa.

Ejemplo 3.1. *Para ilustrar esta situación, consideremos este caso presentado en Bracken (1966). Supongamos que una persona quiere asignar una distribución a priori beta a la variable aleatoria π que representa el voto liberal en una elección futura como la fracción del total de votos. Asumamos que ella encuentra difícil asignar un valor esperado subjetivo al valor de π , pero es capaz de expresar sus juicios asignando una probabilidad de 0.50 a la afirmación $\pi \leq 0.50$ y asignando una probabilidad de 0.75 a la proposición $\pi \leq 0.60$.*

El método CDF también es conocido como *método de intervalo variable*, ya que al experto se le pide cambiar su intervalo en el cual quiera poner una cantidad determinada de probabilidad. El método (O'Hagan, 1998) se inicia con las cotas superior, U ; la inferior, L ; de la variable X y de la moda M , que son elicidadas. El próximo paso es determinar las siguientes cinco probabilidades:

$$\begin{aligned} p_1 &= P(L < X < U) \\ p_2 &= P\left(L < X < \frac{L+M}{2}\right) \\ p_3 &= P\left(\frac{M+U}{2} < X < U\right) \\ p_4 &= P\left(L < X < \frac{L+3M}{4}\right) \\ p_5 &= P\left(\frac{3M+U}{4} < X < U\right) \end{aligned} \tag{3.7}$$

Las probabilidades se preguntan en este orden de tal forma que se evite el anclaje. La selección de los intervalos se hace de tal forma que los sujetos no eliciten probabilidades muy pequeñas.

Las respuestas se convierten en las siguientes probabilidades q_1, \dots, q_6 , de esta forma:

$$\begin{aligned}
 q_1 &= P\left(L \leq X \leq \frac{L+M}{2}\right) = p_2 & (3.8) \\
 q_2 &= P\left(\frac{L+M}{2} \leq X \leq \frac{L+3M}{4}\right) = p_4 - p_2 \\
 q_3 &= P\left(\frac{L+3M}{4} \leq X \leq M\right) = p_1 - p_4 \\
 q_4 &= P\left(M \leq X \leq \frac{3M+U}{4}\right) = 1 - p_1 - p_5 \\
 q_5 &= P\left(\frac{3M+U}{4} \leq X \leq \frac{M+U}{2}\right) = p_5 - p_3 \\
 q_6 &= P\left(\frac{M+U}{2} \leq X \leq U\right) = p_3
 \end{aligned}$$

Note que las probabilidades son más pequeñas alrededor de la moda para producir una mejor elicitación del centro de la distribución.

Perry y Greig (1975) presentan aproximaciones para el cálculo de la desviación típica y de la media de distribuciones subjetivas, a partir de percentiles y de la moda elicitados.

$$\sigma = \frac{p_{95} - p_{05}}{3.25} \quad (3.9)$$

y

$$\mu = \frac{p_{05} + 0.95m + p_{95}}{2.95}, \quad (3.10)$$

donde m es la moda elicitada. Si se posee información sobre la mediana, p_{50} , la media aproximada será:

$$\mu = p_{50} + 0.185(p_{95} + p_{05} - 2p_{50}) \quad (3.11)$$

Aproximaciones similares (pero no iguales) son presentadas por Moder y Rodgers (1968).

Estas aproximaciones son bastante buenas como lo mostramos en una aproximación a la media y la varianza de la distribución Gamma a partir de estos percentiles y la moda de la distribución verdadera.

Tabla 3.5. Aproximaciones para la varianza, la moda, la y media de la distribución gamma con diferentes combinaciones de parámetros

α	β	$\frac{(\sigma_{Aprox} - \sigma_{Teo})}{\sigma_{Teo}}$	$\frac{(Mo_{Aprox} - Mo_{Teo})}{Mo_{Teo}}$	$\frac{(\mu_{Aprox} - \mu_{Teo})}{\mu_{Teo}}$
		Varianza	Moda	Media
1.1	1	-0.0858751203	0.0331019666	0.0004503311
1.1	2	-0.0858751203	0.0331019666	0.0004503311
1.1	5	-0.0858751203	0.0331019666	0.0004503311
1.1	10	-0.0858751203	0.0331019666	0.0004503311
1.1	20	-0.0858751203	0.0331019666	0.0004503311
1.1	50	-0.0858751203	0.0331019666	0.0004503311
2	1	-0.0451876197	0.0252925473	0.0003577095
5	1	-0.0115383128	0.0117741420	0.0001102572
10	1	0.0002343325	0.0061227840	0.0000455299
20	1	0.0062022460	0.0031168010	0.0000201753
50	1	0.0098061020	0.0012595880	0.0000074345
2	2	-0.0451876197	0.0252925473	0.0003577095
5	2	-0.0115383128	0.0117741420	0.0001102572
10	2	0.0002343325	0.0061227840	0.0000455299
20	2	0.0062022460	0.0031168010	0.0000201753
50	2	0.0098061020	0.0012595880	0.0000074345

3.6. Comparación de métodos de elicitación

Varios autores han realizado trabajos en los cuales se comparan diferentes métodos de elicitación. Ludke et al. (1977) comparan cinco procedimientos. O’Leary et al. (2009a) comparan tres métodos de elicitación para regresión logística. El de O’Leary et al. (2009b) comparan cuatro métodos aplicados a modelos logísticos y árboles de regresión. Los métodos son el de Denham y Mongersen, el de Kynn, el de Martin et al. y un método desarrollado para árboles de regresión.

3.7. Técnicas de elaboración de preguntas y entrevistas para la elicitación

La gente que trabaja en el área social donde se recoge información vía entrevistas o mediante cuestionarios autodilenciados, por ejemplo, los que trabajan en el área de mercadeo, reconocen la importancia fundamental de la elaboración de buenas preguntas y preparación de los entrevistadores (Morton-Williams, 1993). Esta es un área difícil y en general no se discute en el caso del proceso de elicitación. Conceptos tales como probabilidad son entendidos en formas diferentes por personas aun con el mismo nivel de conocimiento y generan desconfianza en los procesos de elicitación.

Preguntas que hagan referencia a valores representativos pueden crear imágenes mentales diferentes entre distintos expertos. Términos técnicos de uso diario, tales como la desviación típica, pueden no tener una clara interpretación y confundirse con la desviación media absoluta o con la distancia media de Gini.

3.8. Métodos basados en apuestas

Una de las primeras formas de hacer elicitaciones se basaba en el concepto de *apuesta* que lleva inherente el de utilidad. Bajo el supuesto de utilidad lineal del dinero, es posible crear loterías cuya probabilidad de éxito sea equivalente a la probabilidad de algún evento de interés, basadas en el concepto de *indiferencia*. Harrison et al. (2014) retoman el uso de elicitación vía loterías binarias.

Las apuestas han sido utilizadas en los procesos de elicitación casi exclusivamente entre los estadísticos que hacen elicitación, una aproximación que se les debe a los economistas que trabajan con el concepto de utilidad.

Este proceso está basado en el supuesto de que las creencias pueden ser expresadas en términos de apuestas; esto se hace mucho en la práctica y puede ser utilizado como una forma general de hacer las creencias relativas explícitas (D'Agostini, 2000). Hay condiciones naturales a ser impuestas sobre las apuestas:

- La apuesta debe ser reversible y ninguna puede ser elaborada de tal forma que uno pierda o gane con certeza. Esta condición obliga al sujeto a asignar las apuestas consistentemente con sus creencias.
- La anterior también obliga a aceptar la segunda condición: una vez que ha fijado los odds, debe estar preparado para apostar en cualquier dirección. Esta coherencia juega dos papeles importantes:
 1. Es moral y obliga a la gente a ser honesta,
 2. Es formal y permite que las reglas básicas de la probabilidad sean derivadas como teoremas.

Una probabilidad puede pensarse en términos de las cantidades que se involucren en una apuesta y la porción que el sujeto estuviera dispuesto a arriesgar. Por ejemplo, si usted dice: “Las posibilidades de que el DIM gane el torneo son de uno entre 10,” esto implica que:

$$P(\text{DIM gane el torneo}) = \frac{1}{10}$$

y

$$P(\text{DIM no gane el torneo}) = \frac{9}{10}.$$

Así, es nueve veces más posible que el DIM no gane el torneo a que lo gane, según su opinión. En términos de apostadores se dice que las apuestas están 9 a 1. Si a usted le ofrecen cambiar esta apuesta por una monetaria en la cual se gane 9 pesos con probabilidad $1/10$ y pierda un peso con probabilidad $9/10$, a usted le sería indiferente el cambio. La linealidad de la utilidad exige que esta lotería sea equivalente a una en la cual usted se gane 9.000.000 de pesos con probabilidad $1/10$ y pierda 1.000.000 de pesos con probabilidad $9/10$. Esto, en otras palabras significa que si usted apuesta un peso por el evento “DIM gana el torneo”, si este evento ocurre, usted perdería un peso.

3.8.1. Limitaciones del método de apuestas para elicitación

Las condiciones pueden ser conceptualmente atractivas, por ejemplo, el concepto del *hombre económico*, pero no hay garantía de que dos personas sean capaces de separar sus sentimientos acerca de las apuestas. Existen situaciones en las que las apuestas funcionan como parte integral del sistema, por ejemplo, en finanzas o deportes. Las apuestas a favor o en contra de un resultado particular han sido usadas para determinar las probabilidades de los diferentes eventos. Sin embargo, como se ha mostrado muchas veces, las apuestas son indicadores de las expectativas de los involucrados que, a su vez, pueden ser manipulados por otros actores en el mismo proceso.

3.9. Problemas con la elicitación

Bassett y Lumsdaine (2001) discuten problemas con los procesos de elicitación en los que hacen referencia a aquel de dar respuestas que sean socialmente aceptables. Este tipo de trabajos dan la sensación de que es incorrecto usar información subjetiva debido a los problemas con ellos. No siempre es esta la situación, ya que en la mayoría de los casos los expertos presentan sus creencias con respecto a eventos que son independientes de ellos; no sería así si tratáramos de elicitación de eventos que los afectan directamente, por ejemplo, si quisiéramos determinar la distribución subjetiva sobre la probabilidad de supervivencia de un hijo que se encuentra en un estado crítico, la elicitación, creeríamos, estaría sesgada por los intereses personales del experto. Sabiendo que estos problemas están presentes, queda la responsabilidad por parte de los investigadores de desarrollar procedimientos que los superen.

Un sesgo es una tendencia a tomar en cuenta factores que son irrelevantes para la tarea señalada o de ignorar factores relevantes. Se deben distinguir dos tipos de sesgos: el motivacional y el cognitivo. Los sesgos motivacionales son causados por intereses personales y circunstanciales del experto. Los cognitivos surgen durante el proceso de información de este. El experto puede estar muy influenciado por información que se recuerda fácilmente. Puede hacer supuestos sobre la variable de interés no establecidos por la investigación. También pueden surgir sesgos técnicos como la

incapacidad del sujeto de entender la diferencia entre una distribución muestral y una poblacional.

3.9.1. Peligros en el proceso de elicitación

La habilidad de establecer distribuciones a priori es adquirida, básicamente, mediante un proceso consciente de entrenamiento (Félix, 1976). Un experto puede llegar a determinar cuál o cuáles procedimientos le generan más seguridad en su proceso de establecer su información a priori. Por lo tanto, el estadístico que realiza la elicitación debería estar dispuesto a pasar un buen tiempo conociendo a su experto. A veces, ingenuamente, creemos que un procedimiento funciona bien porque nos produce alguna distribución en forma rápida y decidimos que esta representa bien la información del sujeto.

Muchos autores han señalado problemas que pueden presentarse en los procesos de elicitación. Siu y Kelly en 1998 (Hamada et al., 2008) presentan un resumen de los principales problemas que podemos enfrentar:

1. Problemas con valores cero para las a priori: la asignación de un cero a una densidad no se puede corregir con ningún tipo de dato.
2. Cuando se usa la opinión de los expertos, tenga cuidado con los sesgos generados por la forma en que la gente piensa o siente.
3. Esté alerta con generar a priori demasiado estrechas (varianzas extremadamente pequeñas).
4. Asegúrese de que la información usada en la elicitación de la distribución a priori es relevante para el problema a resolver.
5. Sea cuidadoso cuando elicite distribuciones de parámetros que no son directamente observables.
6. Sea cuidadoso con el conservadurismo. Lo que se quiere ser es realista, no conservador.

3.9.2. Parámetros poblacionales a elicitar

En general, los procesos de elicitación pueden ser más o menos complejos dependiendo del tipo de parámetros a elicitar. Entre los parámetros que son fáciles de elicitar tenemos las medidas de localización (medias y percentiles) y de dificultad media a alta; entre estos tenemos las medidas de dispersión y asociación y los de alta dificultad como parámetros especiales, por ejemplo, el de kurtosis o parámetros en modelos complejos como los de modelos de estructuras de covarianzas.

3.10. Validación del proceso de elicitación

Este es un tema que genera muchas dificultades por las mismas características de la mayoría de los problemas en los que hay necesidad de efectuar la elicitación. Entre estas dificultades está la característica de los eventos, por ejemplo, no son repetitivos o, si lo son, no ocurre en las mismas circunstancias. Otras dificultades surgen con la disponibilidad de los expertos, ya que un proceso puede ser complejo y no es fácil garantizar la disponibilidad de ellos en situaciones diferentes. Existen varios acercamientos a la validación de un proceso de elicitación que pueden hacer referencia a:

- La concordancia entre la distribución obtenida y las creencias del experto. En un proceso de elicitación usualmente el analista termina con la determinación de un modelo matemático, más o menos abstracto; pensemos en un modelo beta para reflejar conocimiento sobre una proporción cuya ecuación puede ser ininteligible para un experto. ¿Cómo validamos este modelo final con las creencias del usuario? Existen varias aproximaciones; una mediante la construcción de gráficos, otra mediante la presentación de medidas de resumen, otra por medio de la generación de muestras hipotéticas poblacionales a partir de simulaciones de valores de las distribuciones construidas.
- Determinación de la calidad del proceso para obtener la información necesaria del experto.

Un área en la que, dadas sus condiciones particulares, se ha podido efectuar un estudio detallado de diferentes técnicas de elicitación y hacer propuestas para mejorar es la meteorología (Murphy y Winkler, 1984; Murphy y Winkler, 1977). Debido a la geografía de los Estados Unidos, es necesario realizar pronósticos diariamente. Esto tiene un gran impacto social y económico, lo cual ha ayudado a generar gran respetabilidad. Los pronosticadores en esta área permanentemente evalúan los diferentes procedimientos usados. Se ha llegado a un consenso generalizado de que los métodos de evaluación subjetivos son los más apropiados dada la complejidad del fenómeno. Habiendo tantos expertos y tantos datos, es fácil establecer quiénes son los expertos que tienen mejores resultados. Meeden (1978) propone procedimientos para comparar expertos basados en el concepto de utilidad.

Una nueva teoría que se propone como alternativa para la utilidad esperada es la *teoría prospectiva*. Argumentan sus autores (Tversky y Kahneman, 1992) que esta teoría explica de mejor forma las violaciones que los sujetos tienen en la teoría de la utilidad esperada. Los elementos de esta nueva teoría son:

- Una función de valor que es cóncava para las ganancias, convexa para las pérdidas y tiene mayor pendiente para las pérdidas que para las ganancias.
- Una transformación no lineal de la escala de probabilidad, la cual sobrevalora las probabilidades pequeñas y subvalora probabilidades moderadas y grandes.

Kahneman y Tversky (1984) comentan que la primera aproximación a la teoría prospectiva está en el trabajo de Daniel Bernoulli en 1738, en el cual él trató de explicar el porqué la gente es, en general, adversa al riesgo y el porqué esta aversión disminuye cuando se incrementa la riqueza. Esta teoría resuelve el problema de la utilidad esperada de Savage, la cual presume un hombre racional y que siempre maximiza su utilidad. En el contexto teórico, uno puede proponer teorías como las anteriores y desarrollar metodologías para elicitación de las probabilidades subjetivas, por ejemplo, apuestas o loterías (Harrison et al., 2014), pero siempre vamos a enfrentar la aversión al riesgo del experto. Otra cuestión que no se resuelve es el del supuesto que el experto para un evento posee internamente una probabilidad numérica y que debemos desarrollar el método para descubrirla. El experto puede más bien tener sentimientos que pueden mejor ser expresados por medio de funciones. Claro que esto vuelve el problema de elicitación más complejo.

3.11. Automatización del proceso de elicitación

Varios investigadores han desarrollado programas computacionales para ayudar a automatizar el proceso de elicitación de ciertos parámetros de interés. Morris et al. (2014) desarrollan un programa llamado MATCH, el cual puede utilizar varias técnicas de elicitación para variables univariadas.

Una línea que se ha evolucionado lentamente es el desarrollo de *software* que permita realizar elicitaciones computarizadas mediante programas que sean adaptables y se ajusten a los valores que el experto proporcione. Un programa de tipo experimental es Elicitor (Low Choy, James y Mengersen, 2009) o Explicit (Grigore et al., 2017) que permite de alguna manera prever el futuro próximo en la elicitación. En *R* está el paquete `expert` desarrollado por Goulet et al. (2009). Las ventajas serían enormes en tareas que son muy repetitivas y que pueden generar errores por cansancio de los participantes. Además, el experto en tiempo real puede ver sus resultados y se le pueden entregar resultados hipotéticos mediante simulación.

Fedde (2009) describe el uso de software especializado, el BET-EF (Bayesian Event Tree for Eruption Forecasting), para elicitación de información de vulcanólogos con el fin de determinar las probabilidades de erupciones en volcanes tales como el Colima en México.

El desarrollo de Inteligencia Artificial (IA) nos abre puertas a procesos de elicitación adaptables a las condiciones que se desarrollen con cada experto y que consideren el tipo de problema sobre el que se trabaja.

3.12. Conclusiones y recomendaciones

Diferentes métodos de elicitación se han presentado a lo largo de este capítulo, en el que se ilustran estrategias para adoptar, dependiendo de los intereses del investigador y del problema que se abordará. Es de resaltar que la aplicación de técnicas de elicitación a expertos requiere, además, de habilidades del elicitador para disminuir los sesgos de información que se puedan obtener debido a factores no controlados tales como los cambios en el estado de ánimo de los expertos (posibles intereses de los mismos en los resultados del estudio) y que no fueron expresados en el proceso de selección de los expertos, entre otros. Estos factores, llevan a centrar la atención del investigador en la correcta elección de la técnica a incorporar, la definición adecuada de los expertos y los mecanismos para reducir el sesgo.

Capítulo 4

Método delphi

El método delphi, cuyo nombre se inspira en el antiguo oráculo de Delphos, fue ideado originalmente a comienzos de los años 50 en el seno del Centro de Investigación Estadounidense RAND Corporation¹ por Olaf Helmer y Theodore J. Gordon, como un instrumento para realizar predicciones sobre un caso de catástrofe nuclear (Cooke, 1991). Desde entonces, ha sido utilizado frecuentemente como sistema para obtener información sobre el futuro (Landeta, 1999). Makridakis, Wheelwright y Hyndman (2008) presentan aplicaciones en el caso de pronósticos.

El procedimiento delphi surge como una respuesta a la necesidad de estructurar una metodología que permitiera extraer información *subjetiva* de personas expertas en un área. La palabra subjetiva para nosotros indica toda aquella información que un individuo posee y que usualmente ha elaborado utilizando de variados procesos sobre un tema; de ninguna manera significa arbitraria. Una forma de llevar a cabo esta cuantificación desarrollada por la RAND Co. en los años cincuenta fue tratar de resolver preguntas sobre armamentismo en la época de la guerra fría. Esta metodología busca información consensuada sobre un tema específico, elaborada en un tiempo relativamente corto y con la participación de varios expertos en un breve. Jolson y Rossow (1971) ilustran el proceso delphi con una aplicación en una corporación norteamericana y dan recomendaciones sobre su puesta en práctica. Ellos

¹Ayyub (2000) afirma: “El desarrollo de métodos estructurados para la obtención de opiniones de expertos fue desarrollado por la corporación RAND (Investigación y Desarrollo) de Santa Mónica, California. La corporación RAND surge como un esfuerzo conjunto de la Fuerza Aérea de los Estados Unidos y el fabricante de aeronaves Douglas en 1946 y que se denominó Proyecto RAND. En su primer año de operación, la RAND predijo que el primer satélite espacial se lanzaría a mediados de 1957. La predicción fue validada con precisión por el lanzamiento del Sputnik ruso el 4 de octubre de 1957. En 1948, RAND se separó de la Douglas como el primer tipo de corporación de expertos. La investigación de la RAND Corp. se puede clasificar en cuatro categorías amplias: (1) metodología, (2) planificación estratégica y táctica, (3) relaciones internacionales, y (4) nuevas tecnologías. Casi todas estas categorías pueden confiar fuertemente en las opiniones de los expertos. En sus primeros días entre la Segunda Guerra Mundial y la guerra de Vietnam, RAND Corp. desarrolló dos métodos para la obtención estructurada de opiniones de expertos: (1) Método delphi, y (2) análisis de escenario”.

presentan tanto el formulario para la captura de la información requerida como los resultados intermedios, mostrando las bondades del procedimiento.

Una ventaja de la metodología delphi es que no requiere que los participantes se conozcan. Esto evita posiciones dominantes o dominadas en las respuestas o los conceptos que se emiten. En algunos casos es posible que los participantes se conozcan, mas no las respuestas dadas en cada una de las rondas (Brown, 1968).

Una de las debilidades de esta metodología radica en la dependencia fundamental en el grupo de expertos. Por esta razón, el grupo coordinador debe estar en contacto permanente con ellos, de tal forma que cualquier inquietud o necesidad sea resuelta de la manera más oportuna. Cuando los expertos son externos, es necesario garantizar su equitativa remuneración económica, realizando previamente un estimativo del tiempo y esfuerzo que este procedimiento requiere.

4.1. Pasos de un delphi

Los pasos del método pueden describirse esquemáticamente así (Landeta, 1999):

1. Definición del problema.
2. Determinación de objetivos (parámetros a elicitar).
3. Conformación del grupo coordinador del delphi.
4. Determinación del grupo potencial de expertos.
5. Elaboración de las herramientas para una ronda (incluye encuestas, documentos, etc.) y ejecución.
6. Análisis y elaboración de resúmenes de esta ronda.
7. Los pasos 5 y 6 se repiten hasta obtener un consenso.
8. Informe final.

Adams (2006) presenta un esquema procedimental que sirve como guía en un proceso de elicitación complejo. Este lo presentamos a continuación.

4.2. Un modelo para el proceso de elicitación

4.2.1. Preliminares

■ Prepare el proceso

- *Confirme la necesidad de la elicitación.* Los procesos de elicitación pueden ser demasiado costosos y desgastantes; por lo tanto, debemos estar convencidos de que es la forma más apropiada de construir las distribuciones a priori. En muchas oportunidades, una buena búsqueda bibliográfica nos permite levantar estas distribuciones realizando metaanálisis. En general, esta última aproximación genera más confianza en los resultados entre los investigadores.
- *Escoja y entrene al facilitador.* Esta persona es la que realmente va a llevar a cabo el proceso y deberá generar no solo la confianza entre los expertos, sino entender todas las sutilezas del proceso.
- *Establezca los objetivos y la aproximación.* La elicitación de parámetros hace parte, generalmente, de un proyecto mayor. Por lo tanto, debemos conocer cuál es la importancia relativa de la elicitación de parámetros dentro de la investigación. Esto nos permite establecer cuáles son los objetivos y determinar inicialmente la aproximación; por ejemplo, elicitaciones cara-cara, vía internet, etc. También, nos ayuda a definir un perfil de los expertos que se necesitan.
- *Determine el presupuesto; tiempo, costo y recursos humanos.*

■ Defina el problema

- Confirme la definición del proyecto.
- Defina las incertidumbres para las cuales se necesitan las a priori.
- Descomponga las incertidumbres si es necesario.
- Obtenga información o evidencia muestral para las probabilidades.

■ Desarrolle el cuestionario

- Defina las preguntas que estén de acuerdo con las incertidumbres y las preguntas directas.
- Defina las preguntas que usan una técnica de escala continua y las respuestas.

■ Selección del experto

La selección de los expertos es un tema mayor, ya que de su selección depende el éxito o fracaso del método. Los expertos se seleccionan no solo por su conocimiento del tema, sino por su deseo de cooperar con el proceso y su disposición

a escuchar otras opiniones y a cambiar las suyas en caso de que las otras sean más convincentes. A pesar de esto sonar lógico, la experiencia muestra que algunos expertos están tan convencidos de sus creencias que se niegan a escuchar otras opiniones que los contradigan. Brown (1968) presenta un caso de predicción de la población mundial en el año 2000 en el que los expertos venían de áreas tales como sociología, demografía y crecimiento poblacional. En el primer cuestionario se les pedía una autocalificación de competencia, se asignaba un número de 1 a 4, siendo 1 una persona muy competente.

- *Determine las categorías de profesionales adecuadas.* Esto depende básicamente de los objetivos y del alcance del proceso.
- *Determine el nivel de experticia requerido.*
- *Establezca el grupo de expertos.* La pregunta ¿Cuántos expertos? es fundamental en esta etapa. Dependiendo de la complejidad del problema y del presupuesto del que se pueden definir el número de estos y su nivel de especialización. En un proceso delphi se recomienda un número de expertos entre 10 y 20, pero no hay reglas absolutas sobre este número. El problema de tener un número muy grande de expertos es logístico. Si el número de expertos es muy pequeño, la variabilidad en las respuestas puede ser muy limitada; aunque en la práctica es común escuchar la ejecución de un delphi en pequeña escala, estos son estudios limitados usualmente por tiempo o por costo.
- *Obtenga el grupo de expertos que va a participar.* Debemos preparar un documento de invitación en el cual se detallan los objetivos de la investigación de tal forma que el experto sea capaz de interesarse y visualizar su papel dentro del proceso. También debe ser claro para el experto cuánto va a recibir como retribución por su participación y cuáles serán sus requerimientos.

4.2.2. Elicitación

- *Establecimiento de citas o la metodología a utilizar si es vía internet.* Dadas las condiciones tecnológicas actuales, realizar un proceso de elicitación vía internet es una forma económica y, en general, disponible en los ambientes académicos. Skipe o Zoom son grandes herramientas que permiten interactuar cara a cara con el experto, permitiendo resolver inquietudes en tiempo real. Sin embargo, el proceso puede contar con varias formas de levantar la información, por ejemplo, mediante cuestionarios para ser desarrollados por los expertos, con discusiones abiertas vía internet de las respuestas dadas por ellos.
- *Presente el facilitador.* El facilitador es la persona que va a ser el medio entre el experto y la representación en papel de sus incertidumbres. Es de importancia extrema tener una persona que entienda el lenguaje, los

prejuicios (heurísticas) y las dificultades que puede tener el experto al expresar su conocimiento y ser capaz de obtener en forma adecuada las incertidumbres. Por lo tanto, el facilitador deberá ser una persona que genere una reacción positiva entre los expertos.

- *Explique el propósito de la elicitación.* Como una fase inicial del delphi se requiere una socialización entre el grupo de expertos de la metodología, haciendo énfasis en las ventajas del procedimiento, por ejemplo, mediante el anonimato es posible dar respuestas que estén en concordancia con la información que un experto tiene, pero que a los otros expertos pudieran no estar de acuerdo. Cabe recordar que toda respuesta es justificada por parte del experto.
 - *Determine los horarios.*
 - *Discuta el proceso de retroalimentación.*
- *Codificación por parte de los expertos*
- Motivación de los expertos y determinación de sesgos.
 - Estructuración del parámetro mediante su clara definición y explicación al experto.
 - Condicionamiento al experto de tal forma que piense fundamentalmente en sus juicios.
 - Codificación de sus juicios probabilísticos.
 - Verificación de las respuestas chequeando su consistencia y viendo si el experto cree en sus resultados.

4.2.3. Análisis de los juicios subjetivos

- *Asignación de valores numéricos de probabilidad de los expertos.* Si contamos con expertos normativos, esta tarea es relativamente fácil de realizar.
- Obtenega estimaciones del rango completo mediante la suavización y escalamiento de las estimaciones de expertos.
 - Convierta los valores suavizados en probabilidades coherentes que sumen uno.
 - Construya la distribución de probabilidad.
- *Asignación de probabilidades numéricas a los valores agregados de los expertos*
- Defina la ponderación apropiada a las estimaciones.
 - Encuentre las estimaciones ponderadas.

- Obtenga valores suavizados de los agregados.
- Convierta los valores suavizados en probabilidades coherentes que sumen uno.
- Construya la distribución de probabilidad.

Protocolos similares al anterior han sido utilizados, discutidos y promovidos por Low Choy, O'Leary y Mengersen (2009) para realizar trabajos aplicados en problemas ecológicos.

4.2.4. Sobre los cuestionarios

La elaboración de los cuestionarios es una etapa crítica del proceso. Estos deben presentar toda la información relevante de tal forma que el experto pueda dar la respuesta que más concuerde con su conocimiento actual del tópico en cuestión. Por lo tanto, deben ser claros y suministrar información que ilustre lo suficiente al experto, pero sin conducirlo a una respuesta predeterminada. La forma de redactar una pregunta puede afectar las respuestas y, en consecuencia, debemos ser extremadamente cuidadosos en las propuestas que sean aceptadas. Los grupos de expertos heterogéneos pueden entender palabras o frases completas en forma diferente. Los cuestionarios tienen al menos dos componentes distinguibles: material de contextualización y preguntas sobre los objetivos propiamente dichos. Un mismo objetivo puede estar sujeto a varias preguntas con el fin de simplificarlo (subdivisión del objetivo) o de validarlo. La parte de contextualización puede tener material adicional que ayude al experto a aclarar definiciones, a compartir datos estadísticos, noticias recientes o resúmenes realizados por el grupo coordinador de rondas previas.

Los cuestionarios se elaboran de manera que faciliten, en la medida en que una investigación de estas características lo permita, la respuesta por parte de los consultados. Preferentemente, las respuestas habrán de poder ser cuantificadas y ponderadas (año de realización de un evento, probabilidad de realización de una hipótesis, valor que alcanzará en el futuro una variable o evento).

Se formularán preguntas relativas al grado de ocurrencia (probabilidad) y de importancia (prioridad), y la fecha de realización de determinados eventos relacionadas con el objeto de estudio: necesidades de información de este, gestión de la información del entorno, evolución de los sistemas, evolución en los costos, transformaciones en tareas, necesidad de formación.

En ocasiones, se recurre a respuestas categorizadas (Sí/No; Mucho/Medio/Poco; Muy de acuerdo/De acuerdo/Indiferente/En desacuerdo/Muy en desacuerdo) y después se tratan las respuestas en términos porcentuales tratando de ubicar a la mayoría de los consultados en una categoría.

La ejecución del cuestionario por parte del experto puede realizarse en forma tal que este no sienta presión o tensión por factores de tipo externo o de tiempo. En todo caso, siempre debe disponer de un integrante del grupo coordinador para que le resuelva cualquier inquietud con respecto al cuestionario o a la información en él plasmada. Este integrante del equipo coordinador no debe generar reacciones negativas; por ejemplo, si el sujeto del equipo coordinador tiene una actitud dominante, puede intimidar a un experto y este sentirse cohibido para dar respuestas apropiadas. Esta ejecución puede tardar desde unos pocos minutos hasta varios días, y depende de la disposición del experto. El tiempo debe manejarse de tal forma que en ningún momento los expertos se sientan cansados u obligados a permanecer por más tiempo del necesario.

Los análisis realizados por el grupo coordinador retroalimentan el grupo de expertos, ya que se resumen los argumentos presentados por estos en los cuestionarios y se presentan medidas y gráficos de tipo estadístico que son el insumo en las siguientes rondas.

Obtención de los principales resultados: los principales estadísticos que se emplearán en el estudio serán medidas de tendencia central y dispersión: media, mediana, moda, máximo, mínimo y desviación típica. Esto nos permite tener una visión de conjunto de los resultados obtenidos en cada una de las preguntas, aunque luego solo se utilicen la media o la mediana como valor para la segunda vuelta. La media y la mediana nos indican la tendencia central de la distribución o conjunto de respuesta de expertos, al igual que la moda. El máximo y el mínimo nos indican las respuestas extremas. La desviación nos señala el grado de dispersión en las respuestas (si más o menos los expertos se hallan en torno a las cifras de la media o no). Los cuartiles, vendrían a ayudar también en la visión del grado de dispersión de las respuestas. El cuartil 1 (Q1), que es igual al percentil 25, sería el valor que deja el 25 % de las respuestas por debajo de ella y el 75 % por encima. El cuartil 3 (Q3), que es igual al percentil 75, sería el valor que deja el 75 % de las respuestas por debajo de ella y el 25 % por encima. Es decir, entre Q1 y Q3 se situaría la mitad central de las respuestas obtenidas.

4.2.5. Ejecución de la segunda vuelta

Como puede observarse en la formulación más académica o teórica del método delphi (Godet, 1996; Landeta, 1999), trabajar buscando la convergencia de los expertos con base en disminuir el espacio intercuartil precisando la mediana supone tener que realizar tres o más vueltas (consultas) a los expertos. Por esto, aunque se han llevado a cabo formulaciones clásicas del método delphi, considerando los objetivos de la investigación y el horizonte temporal relativamente próximo respecto al que encaramos para el estudio prospectivo, optaremos por el siguiente sistema:

Objetivo del lanzamiento y tratamiento de la segunda vuelta

Una segunda vuelta en el estudio se plantea con un doble objetivo:

- a) Remitir y hacer partícipes de la información obtenida a todos aquellos que han colaborado en el estudio con la aportación de su conocimiento y opiniones.
- b) Consolidar y refrendar los resultados obtenidos en la consulta inicial. De hecho, la experiencia indica que las variaciones respecto a los resultados iniciales son mínimas en este tipo de estudios.

Metodología para el lanzamiento y tratamiento de la segunda vuelta

a) Se seleccionan la media o la mediana de las respuestas a las preguntas de la primera ronda y dependerá del tipo de pregunta; si las desviaciones típicas no son excesivas, se utiliza la media.

b) Se solicita a los expertos que indiquen su acuerdo o desacuerdo con dicha media.

c) Se pide a los expertos que no se hallan de acuerdo con la media que argumenten sus razones. ¿Está usted de acuerdo con la media obtenida para el conjunto de los consultados? Sí/No. En caso de que no esté de acuerdo, ¿cuál es el nuevo valor que propone? Si lo juzga necesario, ¿podría justificar dicha respuesta?

d) Se calcula la nueva media o mediana (a los expertos que se muestren de acuerdo se les fijará el valor de la media anterior).

Los valores propuestos y razonados por los expertos que siguen manteniendo valores diferentes de la media sirven para elaborar escenarios alternativos o formular hipótesis de futuro alternativas (el futuro no es único y predeterminado), que serán más valiosas en función de la *calidad* del experto en el tema.

A este respecto, hay que señalar que a veces se ha recurrido a calificar la *competencia* del experto en cada pregunta o bloque de preguntas (también se ha utilizado la autocalificación) y a ponderar las respuestas en función de la calificación del experto. Muy bueno (3), bueno (2), regular (1), no es competente para responder (0). Sin embargo, dado el carácter del estudio no se considera necesaria ésta medida.

Validación del delphi

Dada la forma en que se procede en esta metodología, su proceso de validación puede ser complejo y puede hacerse solo en condiciones relativamente sencillas en las cuales exista un modelo fiable que arroje predicciones y pueda, por lo tanto, realizarse una comparación con los valores generados por el delphi. Otra opción es realizar dos delphi simultáneamente, donde los expertos se aleatorizan. Los resultados finales pueden, por lo tanto, compararse.

En cada etapa se debe realizar un análisis estadístico que cubra diferentes aspectos:

1. Generación de medidas de tipo general como las de localización, dispersión y sesgo. Entre las medidas de localización que se usan se tienen el promedio aritmético y la mediana, siendo esta última la más usada debido a su robustez estadística. También se usan algunos percentiles, tales como el 10, 25, 75 y 90. Entre las medidas de dispersión, en procesos delphi se utiliza el rango intercuartílico. Uno de los objetivos del delphi es la reducción del rango intercuartílico con el transcurrir del proceso.

2. Análisis del grupo. Este hace referencia a la determinación de si existen grupos homogéneos entre los expertos. Si es así, estos grupos deben ser caracterizados. El análisis de Clúster es una técnica multivariada que, a partir del cálculo de distancias entre los expertos, nos permite determinar si existen uno o varios grupos. Es posible, mediante esta técnica, determinar también si existe algún individuo cuyas respuestas puedan considerarse atípicas y, por lo tanto, estaría conformando un grupo de un solo individuo.

4.2.6. Errores en un delphi

Sacman (Ayyub, 2001) encontró que en la aplicación del delphi a menudo se cometen los siguientes errores:

- Se trabaja con cuestionarios pobres. Muchos profesionales de la estadística no recurren a profesionales en el área de elaboración y validación de cuestionarios, como pueden ser los sicometristas y se desarrollan instrumentos deficientes.
- No se atiende a las buenas prácticas estadísticas de muestreo y la reducción de datos. En la selección de un grupo de expertos para un proceso delphi no se tiene una muestra aleatoria de estos ni sus opiniones pueden considerarse representativas desde el punto de vista estadístico.
- No se proporcionan medidas de confiabilidad.
- No se define el alcance, las poblaciones y las limitaciones.
- Involucran métodos de agregación confusos de opiniones de los expertos con predicciones sistemáticas.
- Inhiben la individualidad, promueven la conformidad y penalizan a los disidentes.
- Refuerzan e institucionalizan los primeros hallazgos.
- Puede dar una ilusión exagerada de precisión.
- Se carece de profesionalismo.

El valor del método delphi está en su intento inicial de usarse como una herramienta heurística, no científica, para explorar asuntos desconocidos o vagos que son inaccesibles de otra forma. No es un sustituto para la investigación científica.

Cuestionario para la primera ronda

El cuestionario del delphi es más una guía. Un ejemplo podría ser:

- *La probabilidad de que la economía colombiana crezca en más de un 3% el próximo año es menor que 0.25.*

1.	Estoy fuertemente de acuerdo con la información anterior.
2.	Estoy de acuerdo con la afirmación anterior.
3.	No estoy seguro de la información anterior.
4.	Estoy en desacuerdo.
5.	Estoy fuertemente en desacuerdo.

Justifique su selección:

- *La población en Colombia para 2020 está proyectada en 50.912.429 habitantes. En la actualidad se calcula en 45.508.205 y para 2015 se proyecta en 48.202.617. En la actualidad el consumo per cápita es \$200.000. ¿Usted cree que este consumo per cápita se mantendrá en 2015?*

1.	Estoy fuertemente de acuerdo con la información anterior.
2.	Estoy de acuerdo con la afirmación anterior.
3.	No estoy seguro de la información anterior.
4.	Estoy en desacuerdo.
5.	Estoy fuertemente en desacuerdo.

Justifique su selección.

- *¿Cuáles factores cree usted que promueven el consumo de cemento en el país?*
- *¿Qué porcentaje del cemento vendido se dedica a obras públicas?*

4.3. Combinación de la información de múltiples expertos

En algunas situaciones podemos tener información elicitada a varios expertos. La combinación de las probabilidades proporcionadas por ellos es un asunto difícil (Clemen y Winkler, 1990). Una serie de probabilidades subjetivas se dice que está calibrada si, agrupando todos los eventos con probabilidad cercana a p , la proporción de ellos que realmente ocurre es aproximadamente p (Li y Krantz, 2005). El problema para nosotros es que necesitamos una sola distribución de probabilidad.

¿Cómo combinamos las distribuciones de los expertos? Hay dos métodos populares: la media ponderada y la media geométrica². Si tenemos las distribuciones elicítadas $\xi_1(\theta), \dots, \xi_k(\theta)$, donde k es el número de expertos, entonces la distribución definitiva se calcula como:

$$\xi(\theta) = \sum_{i=1}^k w_i \xi_i(\theta), \quad (4.1)$$

donde $w_i \geq 0$ y $\sum_{i=1}^k w_i = 1$ son los factores de ponderación que miden la calidad de los expertos. Para Clemen y Winkler (1997) es conocido como *Logarithmic Opinion Pool*:

$$\xi(\theta) = K \prod_{i=1}^k \xi_i(\theta)^{w_i}, \quad (4.2)$$

donde $w_i \geq 0$ y son los factores de ponderación que miden la calidad de los expertos y K es una constante de normalización para garantizar que, al final, la función hallada sea una función densidad de probabilidad. Sobre las ponderaciones se han propuesto diferentes procedimientos para su asignación (Winkler, 1968; Van Noortwijk et al., 1992). Por ejemplo:

1. *Ponderaciones iguales.* En este caso suponemos que los expertos son absolutamente comparables y no tenemos alguna información adicional que nos permita discriminar entre ellos.
2. *Ponderaciones proporcionales a un ranqueo de los expertos.* Ranquear es un proceso de comparación relativamente simple y nos permite, dadas las hojas de vida de los diferentes expertos, establecer el orden de importancia de cada uno de ellos. Este procedimiento tiene la desventaja de que no nos permite definir qué tan diferentes son un par de expertos.
3. *Pesos proporcionales a una autoevaluación.* El experto puede, dados algunos criterios establecidos con antelación, autoclasificarse en una escala particular.

²Cooke (1991) considera la familia conocida como la r -norma elemental y definida por una media ponderada o *Linear Opinion Pool* (Clemen y Winkler, 1997)

$$M_r(j) = \left(\sum_{i=1}^E w_i p_{ij}^r \right)^{1/r},$$

y la probabilidad r -norma dada por:

$$P_r(j) = \frac{M_r(j)}{\sum_{k=1}^n M_r(k)},$$

la normalización de las medias. En este caso E es el número de expertos y el i -ésimo experto proporciona un vector de probabilidades p_{i1}, \dots, p_{in} para los eventos A_1, \dots, A_n , una partición del conjunto de interés.

Esto puede mejorarse si a cada experto se le informa quiénes serán parte del grupo de estos, ya que en temas muy especializados los expertos usualmente se conocen y pueden evaluar la calidad de los otros y de ellos mismos con respecto a los demás.

4. Ponderaciones basadas en la experiencia elicitoria de los expertos en las cuales se puedan comparar los resultados reales. Esta sería la situación ideal, ya que, si tenemos los resultados de elicitaciones comparables realizadas por ellos y conocemos los resultados reales, podríamos calificarlos objetivamente. Esto es posible en meteorología, en la que los expertos presentan sus predicciones diariamente y las observaciones climáticas se obtienen casi en forma continua. En nuestro país tenemos las encuestas de expectativas económicas que han sido realizadas durante tantos años y pueden ser usadas para este propósito.
5. Calibración de acuerdo con la información obtenida. Este procedimiento permite evaluar los datos proporcionados por los diferentes expertos mediante resúmenes estadísticos de especificidades de los datos obtenidos. Por ejemplo, en el caso del histograma se podrían ordenar las probabilidades representadas por cada barra del gráfico y determinar las distancias entre los expertos o con respecto a un valor promedio. Estas distancias se usan para ponderar los expertos.

Van Noortwijk et al. (1992) desarrollan un protocolo que puede aplicarse fácilmente. Si se tienen E expertos y se les pide elicitar su información sobre la duración de un objeto, se determina una partición del espacio muestral y a cada experto e se le construye una *verosimilitud*.

$$L_e(\mathbf{p}_e | \mathbf{p}_A) = \prod_{i=1}^m p_{iA}^{w_e \beta n_{ie}} = \prod_{i=1}^m p_{iA}^{w_e \beta n_{ie}}, \quad (4.3)$$

donde:

- $w_e \geq 0$ representa la ponderación del experto e que nos determina nuestra valoración de dicho experto.
- β es determinada como una valoración del grupo completo de expertos.
- n_{ie} es el número virtual de observaciones o tamaño muestral equivalente del experto e para la proporción p_{iA} .

Si se asume independencia de opinión entre los expertos, tenemos la *verosimilitud* total:

$$L(\mathbf{p}_1, \dots, \mathbf{p}_E | \mathbf{p}_A) = \prod_{e=1}^E \left\{ \prod_{i=1}^m p_{iA}^{w_e \beta n_{ie}} \right\} = \prod_{i=1}^m p_{iA}^{\beta \sum_{e=1}^E w_e n_{ie}}, \quad (4.4)$$

donde $\sum_{e=1}^E w_e n_{ie}$ corresponde al número virtual de observaciones atribuidas al grupo de expertos.

Morris (1974) realiza una discusión sobre el valor de un experto, considerando la calidad de la información proporcionada por él en términos del efecto de la decisión que sea tomada a partir de su conocimiento. Esta es una consideración de importancia cuando es necesario recompensar monetariamente a los expertos. Fedde (2009) describe el procedimiento usado para mezclar la información proporcionada por expertos vulcanólogos sobre las probabilidades de diferentes tipos de erupciones. Los expertos son expuestos a un cuestionario en el que deben responder preguntas de control y preguntas de autocalificación. A partir de esta información, se construye un *score* para cada experto.

Lindley (1983) considera la situación en que se tiene la información del experto más la información proporcionada por otros expertos en forma de distribuciones de probabilidad. ¿Cómo combinarlas de tal forma que se pueda incorporar el nivel de confianza que se tenga en la información de estos expertos? Si tenemos K expertos, cada uno con distribución de probabilidad individual ξ_i , $i = 1, \dots, K$, entonces podríamos considerar una mezcla de distribuciones dada por:

$$\xi_{Otros}(\theta) = \sum_{i=1}^K w_i \xi_i(\theta), \quad (4.5)$$

donde los w_i representan los niveles de confianza que se poseen en la información proporcionada por los expertos tal que $0 \leq w_i \leq 1$ y

$$\sum_{i=1}^K w_i = 1 \quad (4.6)$$

Por lo tanto, si la a priori del experto es $\xi(\theta)$, la distribución actualizada será:

$$\xi(\theta|Otros) \propto \xi(\theta)\xi_{Otros}(\theta) \quad (4.7)$$

Lindley y Singpurwalla (1986) presentan una aplicación de la metodología en el campo de la confiabilidad de sistemas, considerando uno que posee n componentes con tasas de fallas desconocidas. Un problema con la metodología propuesta por Lindley es que considera que cada experto tiene el mismo nivel de experticia, cosa que en la práctica no siempre es así.

Leonard (1998) presenta el siguiente procedimiento para mezclar la información elicitada de m expertos: si F_1, F_2, \dots, F_m son las funciones de distribución acumuladas elicitadas para un vector de parámetros $\theta \in \Theta$ y se considera que los expertos tienen una base común de conocimiento; si se considera que las F_i son realizaciones de procesos de Dirichlet independientes en una clase lo suficientemente amplia de FDA con funciones paramétricas respectivas $\{a_i \Lambda(\theta); \theta \in \Theta\}$; las a_1, a_2, \dots, a_m son

parámetros positivos y $\Lambda(\boldsymbol{\theta})$ es la FDA común de las F_i ; por último, la a_i mide la concordancia del i -ésimo experto con el grupo de expertos y Λ es la distribución a priori mezclada, la cual debe ser usada como la base para las inferencias mezcladas para $\boldsymbol{\theta}$.

La inferencia para $\lambda = \Lambda(\boldsymbol{\theta})$, para $\boldsymbol{\theta}$ fijo, y cuando a_1, a_2, \dots, a_m están especificados, puede estar basada en la función de verosimilitud:

$$L(\lambda | \mathbf{a}, \mathbf{F}) = \prod_{i=1}^m \frac{\Gamma(a_i)}{\Gamma(a_i \lambda) \Gamma(a_i(1-\lambda))} F_i^{a_i \lambda - 1} (1 - F_i)^{a_i(1-\lambda) - 1}, \quad (4.8)$$

para $0 < \lambda < 1$ y $F_i = F_i(\boldsymbol{\theta})$.

Esto proporciona una base para la inferencia con respecto a Λ , la cual permite combinar las m distribuciones F_1, F_2, \dots, F_m en una forma coherente. Λ puede ser estimado como:

$$\Lambda^*(\boldsymbol{\theta}) = \frac{\sum_{i=1}^m (a_i + 1) F_i(\boldsymbol{\theta})}{\sum_{i=1}^m (a_i + 1)} \quad (4.9)$$

Ingreso promedio de un taxista. En un trabajo realizado por 12 grupos, cada uno elicitó la información de taxistas sobre el ingreso promedio de uno en un día típico de trabajo, con el fin de determinar el impacto económico que el conjunto de taxis tiene en la economía de la ciudad en un día corriente. La información elicitada fue modelada con una distribución *gamma* (α_i, β_i) , para $i = 1, 2, \dots, 12$. Los valores obtenidos están en la tabla 4.1.

Tabla 4.1. *Parámetros estimados del modelo gamma para cada grupo*

	α_i	β_i
1	3805	27
2	8331	64
3	8047	62
4	6383	44
5	15771	118
6	13621	116
7	8446	64
8	5040	44
9	2285	21
10	2446	22
11	2249	18
12	1784	15

La distribución a priori (asumiendo que los grupos son similares y no tenemos información que nos permita cuantificar la habilidad en su proceso de elicitación) la resumimos como:

$$\xi(\theta) = \frac{1}{12} \sum_{i=1}^{12} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \theta^{\alpha_i-1} \exp(-\beta_i \theta) \quad (4.10)$$

Si utilizamos la distribución calculada a partir de la media geométrica de las a priori obtenemos:

$$\xi(\theta) \propto \theta^{\sum_{i=1}^{12} \alpha_i/12-1} \exp\left(-\theta \sum_{i=1}^{12} \beta_i/12\right). \quad (4.11)$$

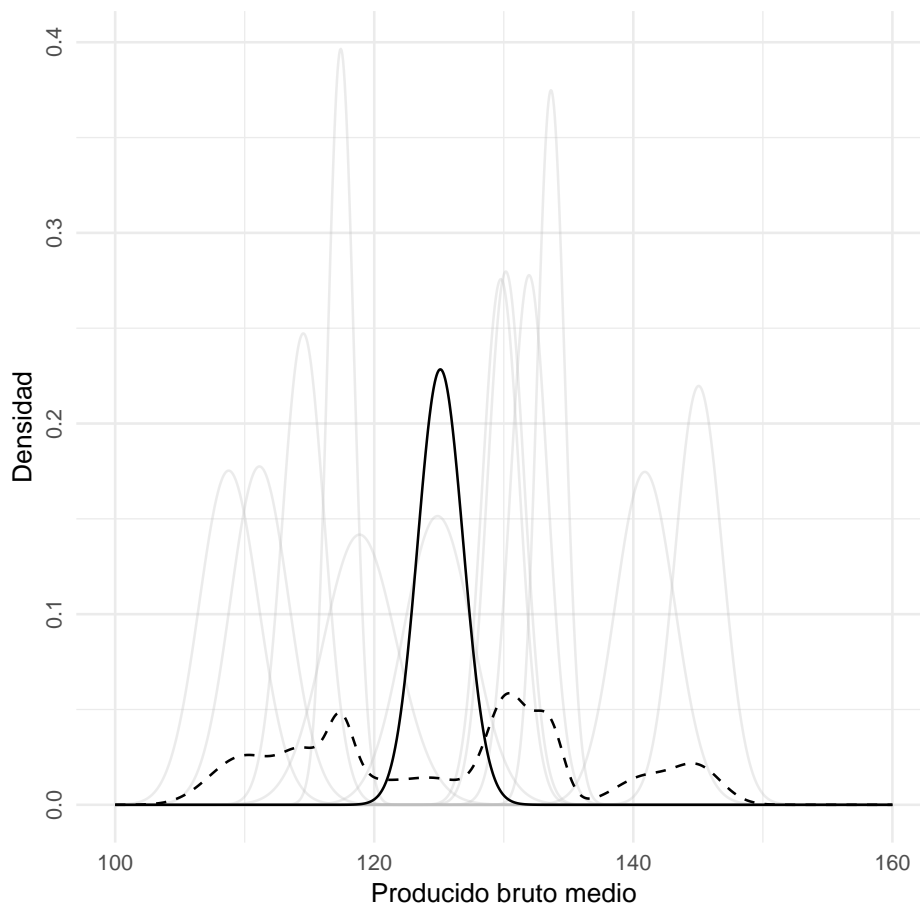


Figura 4.1. *Distribución a priori del ingreso bruto por día en miles de pesos.*

Las funciones elicidadas por los distintos grupos de expertos se representan en color gris. La distribución a priori obtenida a partir de la combinación de estas funciones, mediante una distribución mezclada, se muestra en negro con línea punteada. Por su parte, la línea negra continua representa la distribución a priori calculada utilizando la media geométrica.

Observe la gran variabilidad de la distribución resultante de la mezcla de a priori (línea punteada en la figura 4.1), la cual refleja mucha más incertidumbre que cada una de las distribuciones individuales. Es clara la diferencia entre esta distribución y la resultante cuando usamos la media geométrica (línea negra continua), la cual puede ser considerada una a priori *promedio* en una población de expertos.

4.3.1. Análisis exploratorio de expertos

Cuando se tienen las distribuciones a priori de diferentes expertos, es posible analizar sus datos, de tal forma que se pueda determinar si existen grupos homogéneos dentro de los expertos que permitan realizar una caracterización de estos. Como una ilustración presentamos los resultados obtenidos de un experimento donde se le solicitaba a un conjunto de *expertos* (personas a las que se había preparado en la evaluación de la edad de personas) que construyeran la distribución personal sobre la edad de un sujeto del cual se les presentaba una foto de pasaporte. Luego de obtener las distribuciones a priori, se realizó un análisis de clúster, la técnica multivariable para formar grupos homogéneos y, claramente, se observan dos grupos. Esto puede permitir la realización de dos análisis paralelos, de tal forma que para el mismo problema, estos brinden una visión más amplia sobre las consecuencias de cualquier decisión.

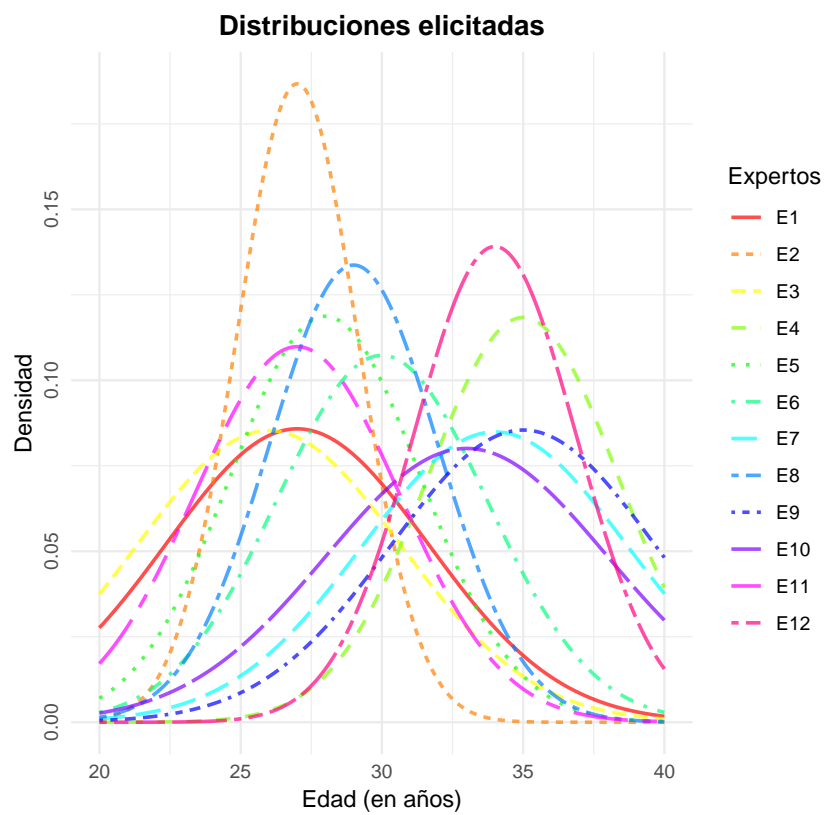


Figura 4.2. Funciones elicidas a 12 sujetos (*expertos*) sobre la edad de un individuo

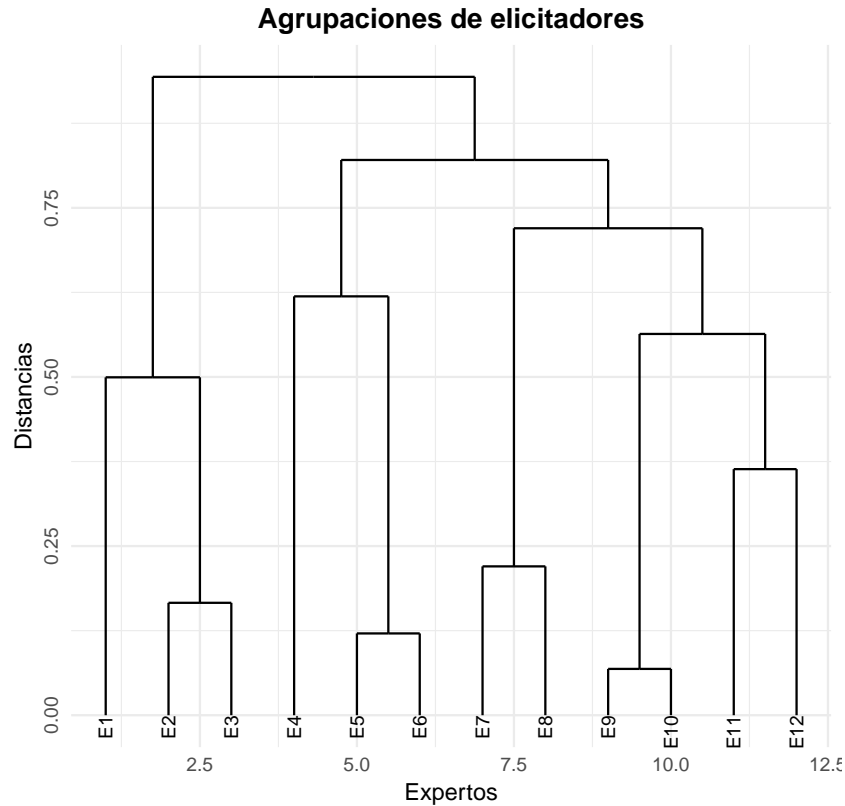


Figura 4.3. *Dendrograma de distribuciones a priori elicidadas*

Si dividimos las a priori elicidadas en tres grupos, el dendrograma muestra que los expertos E1, E2, E5 y E6 forman un primer conjunto con distribuciones similares, lo que indica un alto nivel de consenso entre ellos. Por otro lado, los expertos E3, E4, E7 y E8 conforman un segundo grupo con una estructura distinta, reflejando una percepción diferente del parámetro analizado. Finalmente, los expertos E9, E10, E11 y E12 se agrupan en un tercer conjunto, sugiriendo que sus distribuciones a priori comparten características propias que los diferencian de los otros dos grupos.

Esta segmentación permite simplificar el análisis al trabajar con tres modelos representativos, en lugar de considerar cada experto individualmente. Al observar las diferencias entre los grupos, se pueden identificar patrones en la forma en que los expertos perciben el parámetro, lo que facilita la toma de decisiones y la selección de distribuciones a priori más ajustadas a la realidad del estudio.

4.3.2. Conclusiones y recomendaciones

A pesar de la complejidad en la aplicación del método delphi, esta técnica ha sido ampliamente aplicada en múltiples áreas de las ciencias. La planeación de un proceso de elicitación bajo el método delphi requiere de la aprobación y disposición

previa de los expertos para proporcionar información de calidad en todas las rondas que se lleven a cabo hasta alcanzar consenso acerca de la información elicitada. El análisis exploratorio de expertos es una estrategia clave con fines de acortar el número de rondas del proceso delphi por medio de la obtención de valores iniciales que representen puntos medios de las creencias de los expertos.

Capítulo 5

Aplicaciones

5.1. Elicitación de la media

La media es por excelencia el parámetro más utilizado en estadística para representar una población; aunque, cuando uno habla con usuarios es notable la confusión que se tiene con otros parámetros como la mediana o la moda. Estas confusiones generan problemas con métodos directos de elicitación, asuntos a resolver, considerando métodos robustos ante estas concepciones erróneas. Van Dorp y Kotz (2002) discuten la estimación tradicional de la distribución triangular; sin embargo, en procesos de elicitación esta distribución puede ser de gran ayuda en el inicio, ya que el experto solo debe proporcionar información sobre la moda, el mínimo y el máximo valores plausibles.

Ejemplo: elicitación de la estatura de los hombres adultos jóvenes en Medellín. En este caso, vamos a ilustrar el uso de la técnica de la función de distribución acumulada para elicitar la estatura promedio de los hombres adultos sanos entre 20 y 30 años en Medellín. Para realizar esto, vamos a utilizar la técnica de la muestra hipotética.

Primero determinamos un par de valores, de modo tal que estamos absolutamente seguros de que toda la población de estaturas cae en el intervalo definido por estos valores, que podrían ser 145 cm y 210 cm. Para determinarlos, podemos usar elementos como una cinta métrica, la altura de una puerta típica, etc. Estos elementos nos ayudan a visualizar la población y nos permiten descartar de plano valores imposibles; por ejemplo, un hombre sano no puede medir 100 cm.

Una vez determinado este par de límites, podemos concentrarnos en la mediana: Si tuviéramos una muestra de tamaño 100, ¿cuál sería el valor de la estatura que creyéramos nos dividiría la muestra en dos grupos, uno con los 50 más bajitos y el otro con los más altos? Creemos que 170 cm es ese valor. Este corresponde al percentil 50.

```

# Cargar librería
library(ggplot2)
# Datos
probas <- c(0, 0.10, 0.25, 0.5, 0.75, 0.90, 1)
estatura <- c(145, 163, 167, 170, 175, 178, 210)
datos <- data.frame(estatura, probas)

# Gráfico
ggplot(datos, aes(x = estatura, y = probas)) +
  geom_line(color = "steelblue", size = 1) +
  geom_point(color = "steelblue", size = 2) +
  geom_hline(yintercept = probas, linetype = "dashed", color = "
    gray70") +
  labs(x = "Estatura (cm)", y = "Probabilidad",
       title = "Distribución acumulada de las estaturas\nde los hombres
       de Medellín") +
  theme_minimal(base_size = 13) +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold"),
    axis.text.y = element_text(angle = 90, hjust = 0.5)
  )

```

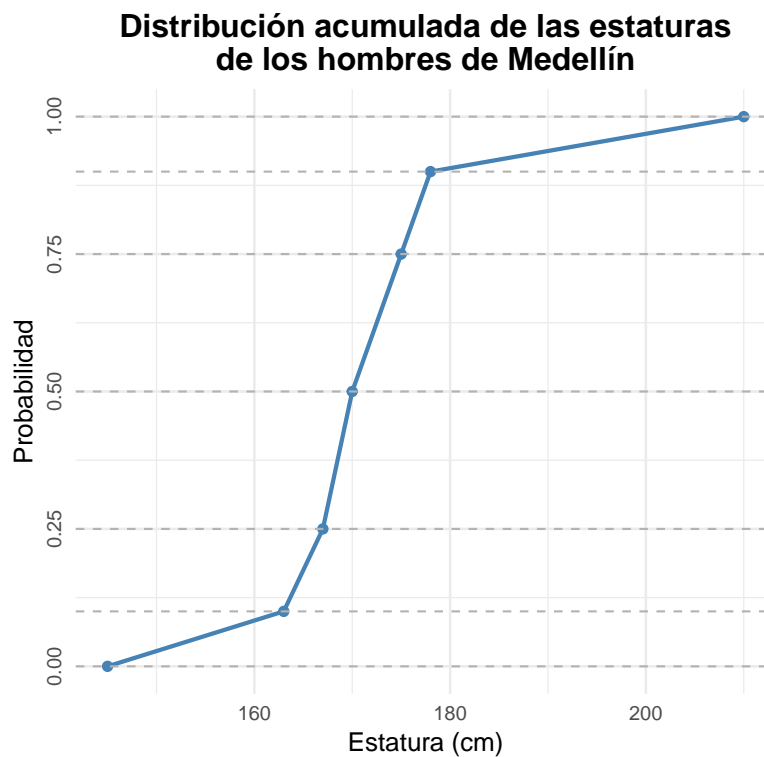


Figura 5.1. *Función de distribución acumulada de la estatura de los hombres en Medellín*

```

# Función auxiliar que dada una probabilidad nos da el percentil
x.s<-function(y.s,x1,x2,p1,p2){
  res<-lm(c(p1,p2)~c(x1,x2))$coefficients
  xx<-(y.s-res[1])/res[2]
  return(xx)
}

# Distribución a priori simulada
dist.apriori.sim<-function(x.s,p.s,Nsim=1000,n.equ=10,funcion='mean')
){
  simulados<-NULL
  y<-sort(runif(Nsim*n.equ))
  for(i in 2:length(x.s)){
    y.temp<-y[y<p.s[i]]
    n<-length(y.temp)
    y<-y[-(1:n)]
    temp<-x.s(y.temp,x.s[i-1],x.s[i],p.s[i-1],p.s[i])
    simulados<-c(simulados,temp)
  }# fin if
  simulados<-matrix(sample(simulados),ncol=n.equ)
  resu<-apply(simulados,1,FUN=funcion)

  plot(density(resu),main='Distribución a priori simulada',ylab='
  Densidad')
  return(resu)
}# fin dist.apriori.sim

res1<-dist.apriori.sim(estatura,probas)

summary(res1)
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 159.5   169.1   171.1   171.3   173.5   182.3

sd(res1)
## [1] 3.327855

```

Distribución a priori simulada

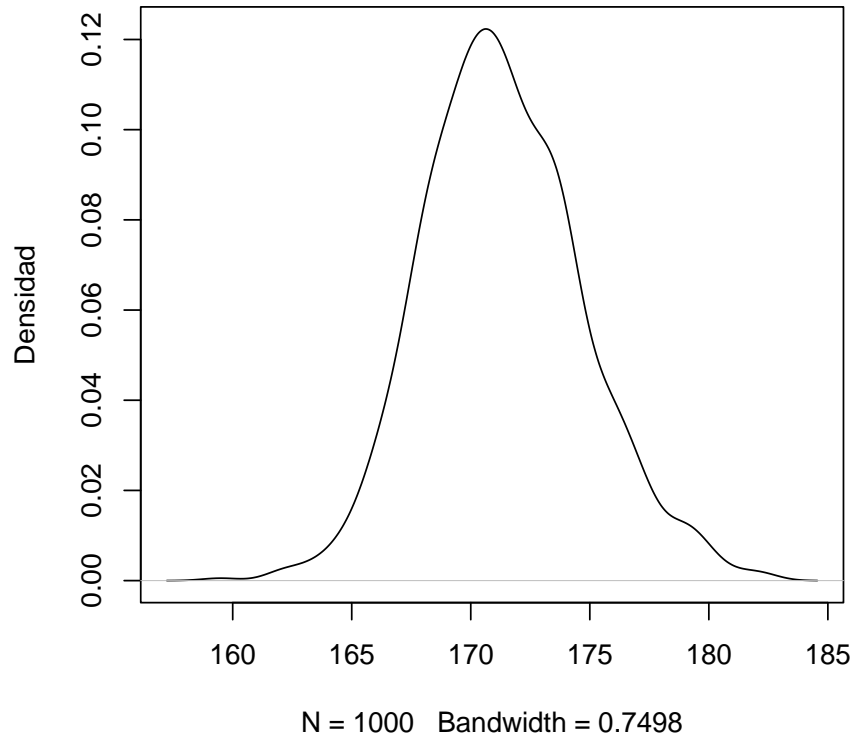


Figura 5.2. *Distribución a priori simulada de la media de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente de nuestro conocimiento a priori igual a 10*

```
res1<-dist.apriori.sim(estatura,probas,n.equ=30)
summary(res1)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  164.4  169.9   171.0   171.1  172.3   181.1
sd(res1)
## [1] 1.874711
```

Distribución a priori simulada

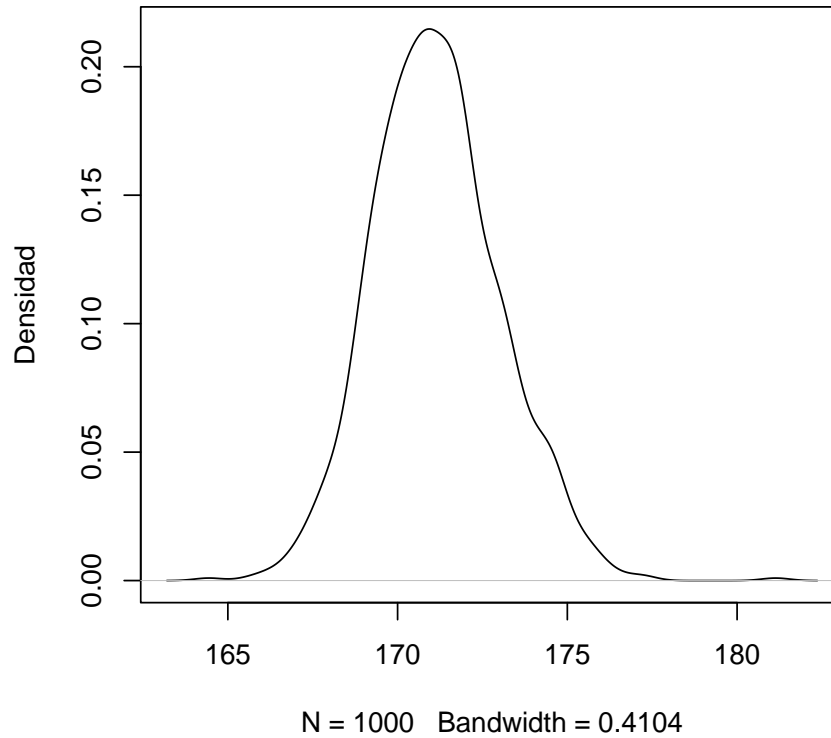


Figura 5.3. *Distribución a priori simulada de la media de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente a nuestro conocimiento a priori igual a 30*

```
res2<-dist.apriori.sim(estatura,probas,funcion='var',n.equ=30)
summary(res2)
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  13.01  75.96  105.96  109.25  138.65  275.17

sd(res2)
## [1] 45.77769
```

Distribución a priori simulada

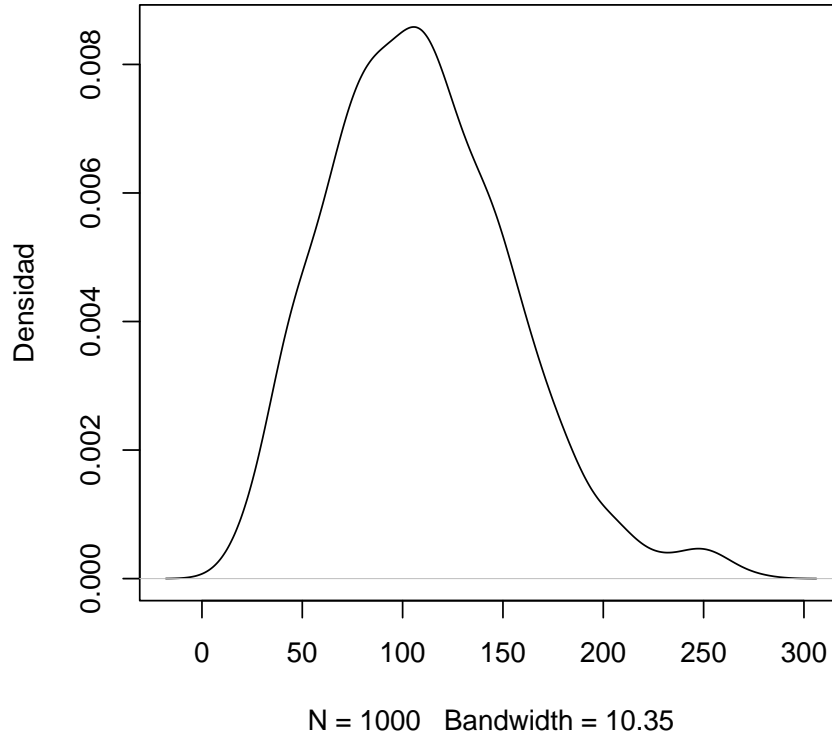


Figura 5.4. *Distribución a priori simulada de la varianza de la estatura de los hombres en Medellín, asumiendo un tamaño muestral equivalente a nuestro conocimiento a priori igual a 30*

5.1.1. Familia conjugada normal: media y precisión desconocidas

Este caso, a pesar de lo simple que puede parecer, muestra la complejidad a la que puede llegar a enfrentarse el estadístico ante la presencia de varios parámetros.

Suponga que X_1, \dots, X_n es una muestra aleatoria de una distribución normal con un valor desconocido de la media μ y un valor desconocido de la precisión τ ($\tau > 0$).

■ **Distribución a priori conjunta de μ y τ**

1. La distribución condicional de μ dado τ es $\mu \sim N(\mu_0, \tau_0\tau)$ donde $\tau_0\tau$ es la precisión, tal que $-\infty < \mu_0 < \infty$ y $\tau_0 > 0$.
2. La distribución marginal de τ es *Gamma* (α_0, β_0) donde $\alpha_0 > 0$ y $\beta_0 > 0$.

■ **Distribución posterior conjunta de μ y τ cuando $\mathbf{X} = \mathbf{x}$**

1. La distribución condicional de μ dado τ es:

$$(\mu | \mathbf{X} = \mathbf{x}) \sim N(\mu_1, \tau_1), \quad (5.1)$$

donde:

$$\begin{aligned} \mu_1 &= \frac{\tau_0\mu_0 + n\bar{x}}{\tau_0 + n} \\ \tau_1 &= (\tau_0 + n)\tau, \end{aligned}$$

y \bar{x} es la media muestral.

2. La distribución marginal de τ es *Gamma* (α_1, β_1) donde:

$$\begin{aligned} \alpha_1 &= \alpha_0 + \frac{n}{2} \\ \beta_1 &= \beta_0 + \frac{1}{2} \sum_{i=1}^n (x_i - \bar{x})^2 + \frac{\tau_0 n (\bar{x} - \mu_0)^2}{2(\tau_0 + n)} \end{aligned}$$

Prueba

Recuerde que:

$$f(x|y) = \frac{f(x, y)}{f(y)} \Rightarrow f(x, y) = f(x|y) f(y) \quad (5.2)$$

Si X_1, \dots, X_n es una muestra aleatoria de una distribución normal con un valor desconocido de la media μ y un valor desconocido de la precisión τ ($\tau > 0$), la verosimilitud será:

$$\begin{aligned} L(\mu, \tau | \text{Datos}) &= \prod_{i=1}^n \frac{\tau^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{\tau}{2} (x_i - \mu)^2\right) \\ &\propto \tau^{n/2} \exp\left(-\frac{\tau}{2} \sum_{i=1}^n (x_i - \mu)^2\right) \end{aligned} \quad (5.3)$$

Ahora:

$$\begin{aligned}
\sum_{i=1}^n (x_i - \mu)^2 &= \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu)^2 & (5.4) \\
&= \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \\
&= (n-1)S^2 + n(\bar{x} - \mu)^2
\end{aligned}$$

$$\begin{aligned}
L(\mu, \tau | \text{Datos}) &\propto \tau^{n/2} \exp\left(-\frac{\tau}{2}((n-1)S^2 + n(\bar{x} - \mu)^2)\right) & (5.5) \\
&\propto \tau^{n/2} \exp\left(-\frac{\tau}{2}(n-1)S^2\right) \exp\left(-\frac{n\tau}{2}(\bar{x} - \mu)^2\right)
\end{aligned}$$

La a priori es:

$$\begin{aligned}
\xi(\mu, \tau) &= \xi(\mu | \tau) \xi(\tau) & (5.6) \\
&\propto \tau^{1/2} \exp\left(-\frac{\tau_0 \tau}{2}(\mu - \mu_0)^2\right) \tau^{\alpha_0 - 1} \exp(-\beta_0 \tau)
\end{aligned}$$

La a posteriori será:

$$\begin{aligned}
\xi(\mu, \tau) &\propto \tau^{n/2} \exp\left(-\frac{\tau}{2}(n-1)S^2\right) \exp\left(-\frac{n\tau}{2}(\bar{x} - \mu)^2\right) & (5.7) \\
&\quad \times \tau^{1/2} \exp\left(-\frac{\tau_0 \tau}{2}(\mu - \mu_0)^2\right) \tau^{\alpha_0 - 1} \exp(-\beta_0 \tau) \\
&\propto \tau^{n/2 + 1/2} \exp\left(-\frac{\tau}{2}[n(\bar{x} - \mu)^2 + \tau_0(\mu - \mu_0)^2]\right) \\
&\quad \times \tau^{\alpha_0 - 1} \exp\left(-\tau\left(\frac{(n-1)S^2}{2} + \beta_0\right)\right)
\end{aligned}$$

Ahora:

$$\begin{aligned}
\left[n(\bar{x} - \mu)^2 + \tau_0(\mu - \mu_0)^2\right] &= n(\mu - \bar{x})^2 + \tau_0(\mu - \mu_0)^2 & (5.8) \\
&= n\mu^2 - 2n\mu\bar{x} + n\bar{x}^2 + \tau_0 - 2\tau_0\mu\mu_0 + \tau_0\mu_0^2 \\
&= (n + \tau_0)\mu^2 - 2\mu(n\bar{x} + \tau_0\mu_0) + n\bar{x}^2 + \tau_0\mu_0^2 \\
&= (n + \tau_0)\left[\mu^2 - 2\mu\frac{(n\bar{x} + \tau_0\mu_0)}{(n + \tau_0)}\right] + n\bar{x}^2 + \tau_0\mu_0^2 \\
&= (n + \tau_0)\left[\mu^2 - 2\mu\frac{(n\bar{x} + \tau_0\mu_0)}{(n + \tau_0)} + \frac{(n\bar{x} + \tau_0\mu_0)^2}{(n + \tau_0)^2}\right] - \\
&\quad \frac{(n\bar{x} + \tau_0\mu_0)^2}{(n + \tau_0)} + n\bar{x}^2 + \tau_0\mu_0^2 \\
&= (n + \tau_0)\left(\mu - \frac{(n\bar{x} + \tau_0\mu_0)}{(n + \tau_0)}\right)^2 - \frac{(n\bar{x} + \tau_0\mu_0)^2}{(n + \tau_0)} + n\bar{x}^2 + \tau_0\mu_0^2
\end{aligned}$$

Luego, la a posteriori queda:

$$\begin{aligned}
\xi(\mu, \tau) &\propto \exp\left(-\frac{\tau}{2}\left[(n+\tau_0)\left(\mu - \frac{(n\bar{x} + \tau_0\mu_0)}{(n+\tau_0)}\right)^2 - \frac{(n\bar{x} + \tau_0\mu_0)^2}{(n+\tau_0)} + n\bar{x}^2 + \tau_0\mu_0^2\right]\right) \\
&\quad \times \tau^{n/2+1/2}\tau^{\alpha_0-1} \exp\left(-\tau\left(\frac{(n-1)S^2}{2} + \beta_0\right)\right) \\
&\propto \exp\left(-\frac{\tau(n+\tau_0)}{2}\left(\mu - \frac{(n\bar{x} + \tau_0\mu_0)}{(n+\tau_0)}\right)^2\right) \\
&\quad \times \exp\left(-\frac{\tau}{2}\left[-\frac{(n\bar{x} + \tau_0\mu_0)^2}{(n+\tau_0)} + n\bar{x}^2 + \tau_0\mu_0^2\right]\right) \\
&\quad \times \tau^{n/2+1/2}\tau^{\alpha_0-1} \exp\left(-\tau\left(\frac{(n-1)S^2}{2} + \beta_0\right)\right)
\end{aligned} \tag{5.9}$$

Ahora:

$$\begin{aligned}
&-\frac{(n\bar{x} + \tau_0\mu_0)^2}{(n+\tau_0)} + n\bar{x}^2 + \tau_0\mu_0^2 \\
&= \frac{-(n\bar{x} + \tau_0\mu_0)^2 + n\bar{x}^2 + \tau_0\mu_0^2}{(n+\tau_0)} \\
&= \frac{-n^2\bar{x}^2 - 2n\bar{x}\tau_0\mu_0 - \tau_0^2\mu_0^2 + n\bar{x}^2 + \tau_0\mu_0^2}{(n+\tau_0)} \\
&= \frac{((n+\tau_0)\tau_0 - \tau_0^2)\mu_0^2 + (n(n+\tau_0) - n^2)\bar{x}^2 - 2n\bar{x}\tau_0\mu_0}{(n+\tau_0)} \\
&= \frac{n\tau_0\mu_0^2 + n\tau_0\bar{x}^2 - 2n\bar{x}\tau_0\mu_0}{(n+\tau_0)} \\
&= \frac{n\tau_0(\mu_0^2 + \bar{x}^2 - 2\bar{x}\mu_0)}{(n+\tau_0)} \\
&= \frac{n\tau_0(\mu_0 - \bar{x})^2}{(n+\tau_0)}
\end{aligned} \tag{5.10}$$

Entonces:

$$\begin{aligned}
 \xi(\mu, \tau) &\propto \exp\left(-\frac{\tau(n+\tau_0)}{2}\left(\mu - \frac{(n\bar{x} + \tau_0\mu_0)}{(n+\tau_0)}\right)^2\right) \\
 &\times \exp\left(-\frac{\tau}{2}\left[\frac{n\tau_0(\mu_0 - \bar{x})^2}{(n+\tau_0)}\right]\right) \\
 &\times \tau^{n/2+1/2}\tau^{\alpha_0-1} \exp\left(-\tau\left(\frac{(n-1)S^2}{2} + \beta_0\right)\right) \\
 &\propto \tau^{1/2} \exp\left(-\frac{\tau(n+\tau_0)}{2}\left(\mu - \frac{(n\bar{x} + \tau_0\mu_0)}{(n+\tau_0)}\right)^2\right) \\
 &\times \tau^{\alpha_0+n/2-1} \exp\left(-\tau\left(\frac{(n-1)S^2}{2} + \beta_0 + \frac{n\tau_0(\mu_0 - \bar{x})^2}{2(n+\tau_0)}\right)\right)
 \end{aligned} \tag{5.11}$$

Con esto, queda demostrado el resultado.

```

# Determinación de los parámetros de una normal a partir de # una
# muestra de tamaño n representada en intervalos

estimaNormal <- function(parametros,frecu=frecu,limites=limites){
  media<-parametros[1]
  vari<-parametros[2]
  dev.tip<-sqrt(vari)
  proba.esti<-frecu/sum(frecu)
  error<-0.0
  for(i in 1:length(frecu)){
    error<-error+(pnorm(limites[i+1],mean=media,sd=dev.tip)-pnorm(
      limites[i],mean=media,sd=dev.tip)-proba.esti[i])^2
  }
  return(error)
}

limites<-c(140,160,170,180,190)
frecus<-c(10,40,45,5)

#estimaNormal(c(170,20),frecu=frecus,limites=limites)
(resu<-optim(c(170,25),estimaNormal,method='L-BFGS-B',
lower=c(160,0.01),upper=c(185,100),
limites=limites,frecu=frecus))

```

```

## $par
## [1] 170.25379 48.22593

## $value
## [1] 0.0021774

## $counts
## function gradient
##      13      13

## $convergence
## [1] 0

## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"

```

```

# Nivel de seguridad en nuestro conocimiento n.seg=5

param.opt<-resu$par

calcula.media.y.precision<-function(xx) c(mean(xx),1/var(xx))

n.seg<-5
Nsim<-1000
tempo<-matrix(apply(matrix(rnorm(n.seg*Nsim,mean=param.opt[1],
sd=sqrt(param.opt[2])),ncol=n.seg),1,calcula.media.y.precision),
ncol=2,byrow=T)

library(MASS)
fitdistr(tempo[,2], 'gamma')
##      shape      rate
##  1.6143577 42.3748268
## ( 0.0660815) ( 2.0300159)
## Warning message:
## In densfun(x, parm[1], parm[2], ...) : NaNs produced

fitdistr(tempo[,1], 'normal')
##      mean      sd
## 170.29241357 2.96915833
## ( 0.09389303) ( 0.06639240)

```

```

histograma<-function(frecu,limi){
  probas<-frecu/sum(frecu)
  altura<-probas/(limi[-1]-limi[-length(limi)])
  cotax<-c(0.90*min(limi),1.1*max(limi))
  cotay<-c(0.0,1.2*max(altura))
  plot(cotax,cotay,type='n')
  for(i in 1:length(frecu)){
    #altura<-probas[i]/(limi[i+1]-limi[i])
    polygon(c(limi[i],limi[i],limi[i+1],limi[i+1]),c(0,altura[i],
      altura[i],0),col='gray')
  }
}

histograma(frecus,limites)

points(xx<-seq(min(limites),max(limites),length=100),
  dnorm(xx,mean=param.opt[1],sd=sqrt(param.opt[2])),type='l')

```

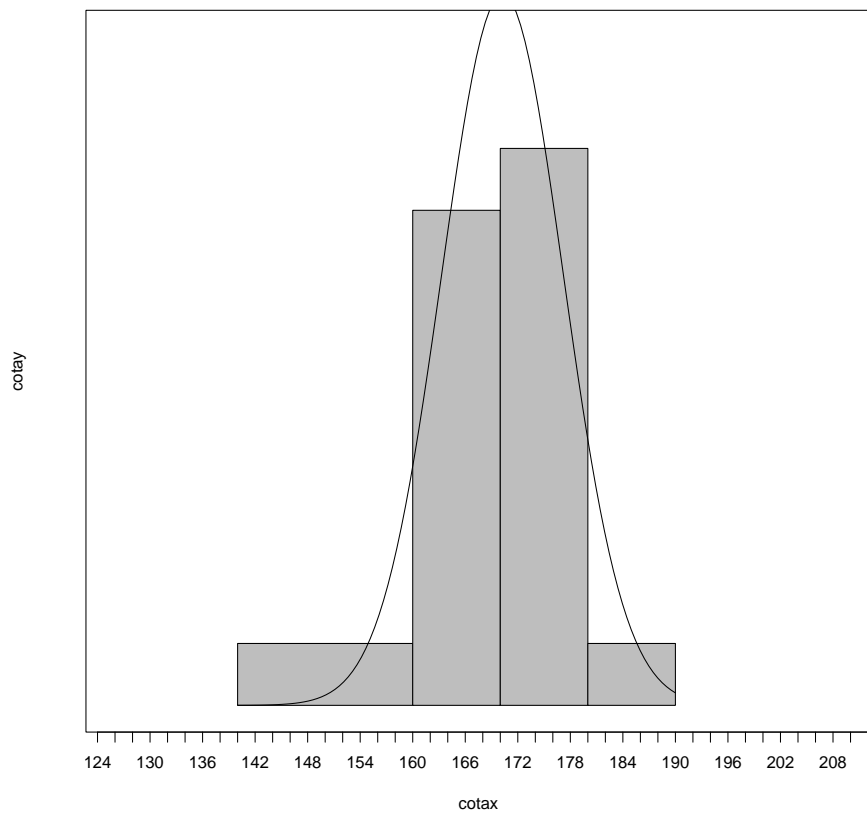


Figura 5.5. Histograma y curva normal ajustada a partir de la elicitation de los parámetros del modelo.

Generemos muestras simuladas de la distribución normal conjugada.

Algoritmo

1. Fije el número de pares (μ, τ) a generar, digamos, $Nsim$
2. Genere $Nsim$ valores τ de una $Gamma(\alpha_1, \beta_1)$
3. Para cada valor de τ genere un μ de una $Normal(\mu_1; \tau \cdot \tau_1)$

Programa en R

```
auxi<-function(tau,mu1,tau1) rnorm(1,mean=mu1,sd=1/sqrt(tau*tau1))

genera.muestras<-function(mu1,tau1,alfa1,beta1,Nsim=10000){
  taus<-rgamma(Nsim,mu1,rate=beta1)
  medias<-sapply(taus,auxi,mu1,tau1)
  return(cbind(medias,taus))
}
```

```
resu<-genera.muestras(170.2924, 2.9691, 1.6143, 42.3748)
summary(resu)
##      medias      taus
##  Min.   :169.3   Min.   :2.867
##  1st Qu.:170.1   1st Qu.:3.809
##  Median :170.3   Median :4.012
##  Mean   :170.3   Mean   :4.022
##  3rd Qu.:170.5   3rd Qu.:4.226
##  Max.   :171.4   Max.   :5.339

quantile(resu[,1],probs=c(0.025,0.5,0.975))
##      2.5%      50%      97.5%
## 169.733 170.288 170.863

quantile(resu[,2],probs=c(0.025,0.5,0.975))
##      2.5%      50%      97.5%
## 3.442548 4.012304 4.659154
```

5.2. Métodos indirectos

Los métodos directos son populares entre los teóricos y fueron propuestos en los comienzos de la teorización de la probabilidad por Ramsey y De Finetti en los años 30 (Ayyub, 2000). En cambio, los métodos indirectos tratan de tomar ventaja de situaciones mentales del sujeto que son más intuitivas.

Los métodos indirectos, en los cuales se recurre a situaciones relacionadas con el proceso de observar eventos que pudieran ser resultantes del fenómeno bajo consideración, son atractivos por parte del elicitador, ya que nos generan resultados que son consistentes con la realidad subjetiva del experto.

5.2.1. Apuestas

Si estamos interesados en determinar la probabilidad que un sujeto tenga sobre la ocurrencia de un evento A , se puede proceder de una forma como:

El sujeto recibe un premio de x pesos si A ocurre.

Alternativamente, se le ofrece escoger una lotería con la que puede ganar los x pesos con k boletas de un número total de M boletas. Se debe buscar el número x_I tal que para la persona sea *indiferente* entre escoger la situación en la que recibe el premio mediante la ocurrencia del evento A o mediante la lotería. Cuando se obtenga ese punto de indiferencia, entonces se iguala.

$$\text{Probabilidad}(A) = \frac{x_I}{M} \quad (5.12)$$

5.2.2. Método de verosimilitudes relativas

En esta aproximación se pregunta al experto sobre la posibilidad de ocurrencia de un evento con respecto a otro. Una forma: ¿qué tan posible es que un valor sea el verdadero con respecto a otro? Por ejemplo, en el caso de la edad de una persona, es posible presentarle al experto la foto del sujeto y preguntarle sobre cuántas veces es más probable que el sujeto tenga una edad de 40 a 45 años.

5.2.3. Previsión de muestras futuras

Este es uno de los métodos más fáciles para aplicar elicitaciones en el caso en el que se disponen modelos hipotéticos y el objetivo final es la elicitación de las distribuciones a priori conjuntas de estos parámetros. Casos como el modelo lineal se pueden aproximar mediante esta técnica. La aplicación es muy demandante por parte de las personas involucradas en el proceso, ya que exige una interacción permanente entre el experto y el equipo técnico. Conceptualmente, es fácil de entender; solo se necesita imaginar valores de muestras hipotéticas plausibles que se obtengan a partir de la población de referencia. Este es el procedimiento que recomendamos.

Muestras hipotéticas futuras (HFS)

Se le pide al sujeto que imagine obtener una muestra real de sujetos de la población de interés. Aquí hay un problema que se debe resolver previamente y es el tamaño de esta muestra, ya que este representa realmente el nivel de conocimiento

sobre el parámetro que el sujeto tiene. El método HFS le pide al experto que actualice su estimación de la proporción a la luz de una muestra hipotética obtenida de la población; por ejemplo, si se extrajo una muestra de 30 elementos de los cuales 10 tienen la característica. Este proceso se repite tantas veces como sea necesario para obtener la distribución a priori.

Kadane y Wolfson (1998) recomiendan usar distribuciones predictivas en el proceso de elicitación en lugar de cuestiones estructurales (cuando se elicitan los parámetros de una distribución) para evitar errores que surgen en la elicitación, los cuales están asociados con procesos mentales conocidos como heurística (Tversky y Kahneman, 1974) como:

- *Representatividad*: los expertos determinan la probabilidad en la medida en que ellos aprecian una proposición conocida como representativa de otra nueva. De esta forma los estereotipos pueden afectar la determinación de probabilidades.
- *Disponibilidad*: los expertos encadenan sus probabilidades a la frecuencia con que ellos pueden recordar un evento.
- *Ajuste y anclaje*: los juicios son anclados a algún valor inicial y ajustado, a menudo en forma insuficiente.
- *Sobreconfianza*: esto dificulta la determinación de las colas de la distribución; esto hace parte del conservadurismo. El término *conservadurismo* (Chesley, 1975) es usado para señalar que los sujetos revisan sus probabilidades subjetivas a la luz de la evidencia en menor forma que la prescrita por el teorema de Bayes. Se consideran tres razones básicas para el conservadurismo:
 - Teoría de mala agregación: considera que los individuos no son capaces de calcular o agregar en la forma predicha por el teorema de Bayes.
 - Teoría de la percepción errónea: involucra subestimar o sobreestimar la verosimilitud y así reducir su impacto en la distribución a priori.
 - Sesgo de respuesta: los sujetos son reacios a reportar probabilidades extremas.

Bottomley et al. (2000) proponen dos métodos: el de rating directo y el de distribución de puntos, que pueden compararse con los métodos de muestras hipotéticas combinados con otros como el de intervalos fijos.

Elicitación mediante muestras hipotéticas futuras para una proporción

Vamos a considerar la determinación de la distribución subjetiva que posee un experto acerca del número de estudiantes mujeres de la universidad que ejercen la prostitución. Este es un tópico difícil y delicado, el cual se sabe que mediante encuestas puede no resolverse confiablemente. En una situación como esta, la misma

definición de *experto* puede ser difícil y podríamos considerar varios tipos: técnicos (médicos, trabajadores sociales, etc.) o no técnicos (algunas estudiantes pueden conocer muy bien el comportamiento de sus compañeras). Un problema presente es el sesgo de conocimiento: pertenecer a ciertos grupos sociales puede hacernos creer que se es experto, cuando no es así. Por ejemplo, un médico que esté involucrado en el tratamiento profiláctico de prostitutas puede sobrestimar el número de estas.

Una vez tengamos nuestro experto, podemos desarrollar una conversación como la siguiente:

Entrevistador: Como lo reconoce la comunidad académica en esta universidad, usted es considerado un experto en prostitución femenina entre las estudiantes aquí.

Experto: Gracias. Conozco algo, pero no me considero como un experto. He trabajado en programas de prevención de enfermedades de transmisión sexual con ambos géneros.

Entrevistador: Para nosotros es un honor que nos ayude con el problema de la determinación del porcentaje de estudiantes mujeres que ejercen la prostitución como la principal fuente de sus ingresos; queremos decir que no consideraremos aquellas estudiantes que ocasionalmente obtienen algún dinero de esta forma. Observamos que la definición es poco clara y es un problema metodológico que podría mostrar nuestra poca preparación en el proceso, ya que esta definición debió estar claramente establecida antes de la entrevista. En este punto es posible preguntarle al experto qué entiende por prostitución.

Experto: ¡Difícil! Yo no estoy seguro de que pueda ayudar, pero trataré. Podría considerar una prostituta aquella chica que la ejerce con el propósito de obtener regalos, etc.

Entrevistador: No. Nos vamos a limitar al caso donde la persona obtiene sus medios de supervivencia (cubre todos sus gastos) del ejercicio de la prostitución.

Experto: Ya veo...

Entrevistador: Empecemos. No sienta que tiene que responder de una forma única y cerrada a las preguntas. Si tiene dudas, me las comenta. Si sacáramos una muestra de 100 estudiantes (mujeres) aquí en la universidad, ¿cuántas creería encontrar que cumplan con la definición de prostituta que hemos dado antes?

Experto: No sé... ¿tal vez una o dos? ¡Es difícil!

Entrevistador: No se preocupe, apenas estamos empezando y todas estas cifras las podemos tomar como parte del calentamiento. ¿Si aumentáramos la muestra a

1000, creería, de acuerdo con los números que me dio antes, esperaría que el número de prostitutas estuviera entre 10 y 20?

Experto: Creo que 20 sería mucho. Me voy con 10, máximo.

Entrevistador: O sea, teniendo alrededor de 2500 estudiantes mujeres, ¿cree que podría haber alrededor de 25 prostitutas en la universidad?

Experto: Sí, es posible.

Entrevistador: Si cambia de opinión, por favor me lo hace saber. De todas formas, en nuestra próxima reunión, y después de que haya pensado mejor las respuestas que me acaba de dar, podría darme otros valores.

Experto: Está bien. Este proceso es algo tedioso...

Entrevistador: Sí, pero es importante. ¿Se imagina preguntarle a una estudiante si es prostituta? Ahora, vamos a seguir con el proceso. En esa muestra hipotética de 1000 estudiantes (en la cual usted conoce la verdadera situación de una estudiante), ¿cuántas prostitutas, a lo máximo, podría considerar un valor posible, por ejemplo, 100 sería posible?

Experto: ¡De ninguna forma! Ese número es demasiado alto. A lo más, 20 sería un valor que aceptaría

Entrevistador: Y ¿qué hay de 30?

Experto: Si me dice que 30, en las 2500 estudiantes tendríamos 75..., es posible.

Entrevistador: Y ¿qué pasaría con 40? ¿Sería posible?

Experto: No lo creo, ya sería un número excesivo.

Entrevistador: Ahora pensemos en la muestra de tamaño 1000 nuevamente. ¿Esperaría que en esta muestra no hubiera una sola prostituta?

Experto: No. Al menos 5. Si me dice un número menor no se lo creería.

Vamos a suponer que el proceso se acabó con los valores: 5, 10 y 30. Estos números, divididos por 1000 nos dan unas proporciones aproximadas de 0.005, 0.010 y 0.030.

Como una aproximación de la a priori a una distribución $Beta(\alpha, \beta)$, podemos hacer.

$$\begin{aligned}E(\pi) &= 0.010 \\P(\pi \geq 0.030) &= 0.05 \\P(\pi \leq 0.005) &= 0.05\end{aligned}$$

Si ajustamos una distribución paramétrica como la beta, estos valores nos permiten hallar unos valores aproximados de $\alpha = 0.01768345$ y $\beta = 2.05623039$.

```
ajuste.beta<-function(teta, valores=valores){
  alfa<-teta[1]
  beta<-teta[2]
  cuantil0.05<-valores[1]
  cuantil0.95<-valores[3]
  media<-valores[2]
  cuant1.teo<-qbeta(0.05, alfa, beta)
  cuant2.teo<-qbeta(0.95, alfa, beta)
  media.teo<-alfa/(alfa+beta)
  res<-(cuantil0.05-cuant1.teo)^2
  +(cuantil0.95-cuant2.teo)^2
  +(media-media.teo)^2
  return(res)
}

valores<-c(0.005,0.01,0.02)

optim(c(1,1), ajuste.beta, method="L-BFGS-B",
      lower=c(1,1)/1000000, upper=c(10,10),
      valores=valores)

## [1] 0.5716812 1.0503667
```

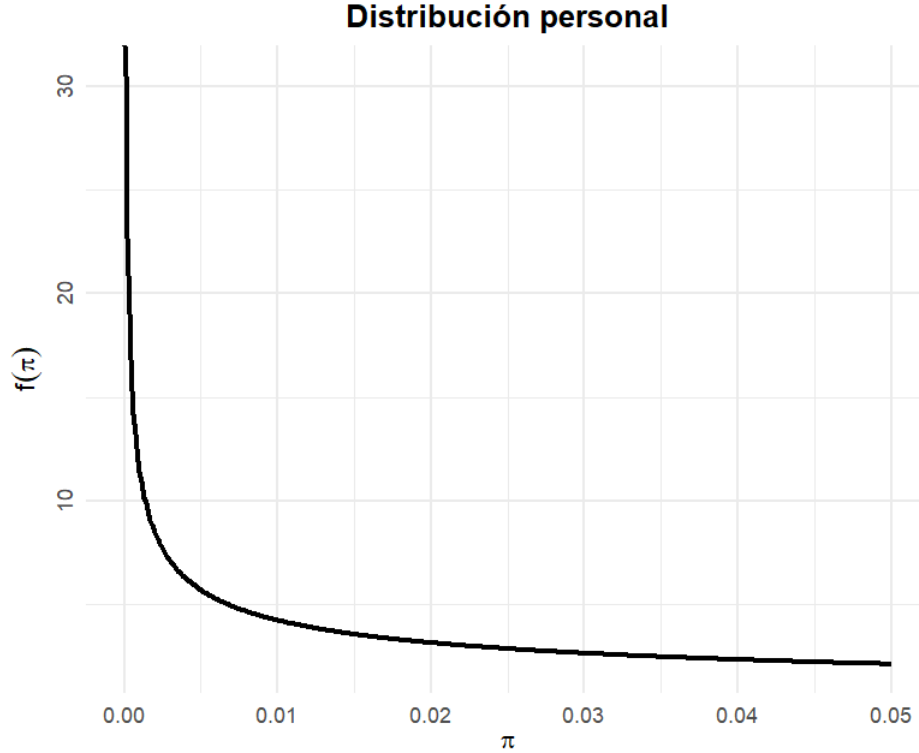


Figura 5.6. *Distribución elicitada paramétrica de la proporción del número de estudiantes que ejercen la prostitución como medio de supervivencia*

Elicitación de la distribución a priori conjugada para el parámetro de la Bernoulli

La distribución beta puede ser asimétrica y la moda puede ser el mejor parámetro a elicitar que un momento como la media (Gavasakar, 1988). El método de Chaloner y Duncan (1983) es conocido como el PM (la moda posterior) y los pasos son:

1. Determine un n hipotético.
2. Estime m (la moda) como el valor más probable.
3. Grafique la distribución $binomial(n, m/n)$. Observe que la probabilidad de éxito corresponde al valor más probable.
4. Calcule las pendientes d_l y d_u como sigue:

$$d_l = \frac{p(m-1)}{p(m)}, \tag{5.13}$$

y

$$d_u = \frac{p(m+1)}{p(m)}, \quad (5.14)$$

donde $p()$ es la probabilidad predictiva del sujeto. Estas cantidades deben satisfacer la siguiente desigualdad:

$$d_l d_u > \frac{m(n-m)}{(m+1)(n-m+1)} \quad (5.15)$$

5. α y β se determinan resolviendo:

$$d_l = \frac{(n-m)(m+\alpha)}{(n-m+1)(m+\alpha-1)}, \quad (5.16)$$

y

$$d_u = \frac{m(n-m+\beta)}{(m+1)(n-m+\beta-1)} \quad (5.17)$$

6. Luego, viene un proceso de calibración. Si se denotan por α_1 y β_1 los parámetros determinados en el paso anterior, entonces, se procede así:

a) Se calcula la moda elicitada a priori como:

$$\gamma = \frac{\alpha_1 - 1}{\alpha_1 + \beta_1 - 2} \quad (5.18)$$

b) Calcule el intervalo de probabilidad más corto del 50% y pregúntele al experto si está de acuerdo con este intervalo.

- Si lo considera muy largo, haga $h = -1$
- Si lo considera muy corto, haga $h = 1$
- Si lo considera adecuado, haga $h = 0$

c) Reajuste los parámetros a priori como:

$$\alpha_{i+1} = 1 + 2^k(\alpha_i - 1), \quad (5.19)$$

y

$$\beta_{i+1} = 1 + 2^k(\beta_i - 1) \quad (5.20)$$

Estos valores nuevos no cambian la moda.

7. Si h no es cero, estos pasos se repiten hasta que h cambie de signo. Aquí se puede hacer un reajuste.

Otra forma de elicitación de los parámetros de la beta para proporciones

- Determine la probabilidad r de que un elemento sacado al azar sea un éxito. Esta probabilidad será considerada la media de la beta:

$$r = \frac{\alpha}{\alpha + \beta} \quad (5.21)$$

- Dada la información de que el primer elemento sea un éxito, determine la probabilidad, r^+ , de que el segundo elemento seleccionado al azar sea otro éxito. La regla dice que la densidad actualizada es una $beta(\alpha + 1, \beta)$, así que:

$$r^+ = \frac{\alpha + 1}{\alpha + \beta + 1} \quad (5.22)$$

- Resuelva simultáneamente:

$$\begin{aligned} \alpha &= \frac{r(1 - r^+)}{r^+ - r} \\ \beta &= \frac{(1 - r)(1 - r^+)}{r^+ - r} \end{aligned} \quad (5.23)$$

- Chequee consistencia: pregunte por la probabilidad de que el segundo artículo sea un éxito, dado que el primero fue un fracaso, diga r^- . Utilizando los valores de α y β , calcule:

$$r^- = \frac{\alpha}{\alpha + \beta + 1}, \quad (5.24)$$

y compruebe si los valores elicitados concuerdan.

- Si el resultado no es satisfactorio, se pueden ajustar los valores de r y r^+ o r^- hasta obtener un resultado consistente.

Elicitación de la distribución a priori conjugada para el parámetro de la Poisson

- Suponga que deseamos estudiar el número de goles marcados por los equipos locales en el torneo profesional colombiano. Asumimos que el número de goles marcados por el equipo local se puede modelar mediante la Poisson.
- El parámetro λ en la Poisson es la media. ¿Cómo escogemos la $gamma(\alpha, \beta)$ que represente adecuadamente nuestro conocimiento del problema?

- Vamos a presentar una aproximación usando una forma predictiva.

Probabilidad	0	1	2	3	4	5	6 o más
$\lambda^x \exp(-\lambda)/x!$	π_0	π_1	π_2	π_3	π_4	π_5	π_{6+}

- Si miramos con cuidado el problema, lo convertimos en multinomial.
- Debemos determinar el vector $(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_{6+})$. ¿Cómo lo hacemos?
- Le decimos al experto que nos responda algo como esto: Si usted observara 1000 partidos de fútbol, ¿en cuántos esperaría que el equipo local no hiciera goles?, ¿un gol?, ¿dos goles?, ¿tres?, ¿cuatro?, ¿cinco?, ¿seis o más goles? Esto nos da una tabla como la siguiente:

Goles	0	1	2	3	4	5	6 ó más.
Número de juegos	n_0	n_1	n_2	n_3	n_4	n_5	n_{6+}

$$\sum_i n_i = 1000$$

Resultado del experto:

Tabla 5.1. *Distribución de frecuencias para el número de goles marcados por los equipos locales en el torneo de fútbol profesional colombiano*

Goles	0	1	2	3	4	5	6 o más
Número de juegos	170	250	300	180	60	35	5

- Generamos N muestras de tamaño 1000 de una multinomial con probabilidades:

$$(170/1000, 250/1000, 300/1000, 180/1000, 60/1000; 35/1000, 5/1000)$$

- Para cada muestra multinomial, calculamos la probabilidad de cada celda, o sea, dividimos cada muestra por 1000. Digamos:

$$(\pi_0^j, \pi_1^j, \pi_2^j, \dots, \pi_{6+}^j)$$

- Usando estas probabilidades, calculamos la media de la distribución de Poisson, teniendo en cuenta que la última celda corresponde a un truncamiento.

$$\lambda_j = \sum_{i=0}^{\infty} i \cdot \pi_i^j \approx \sum_{i=0}^6 i \cdot \pi_i^j \quad (5.25)$$

La aproximación siempre es por debajo del verdadero valor, ya que se reemplazan todos los valores mayores que 6 por 6.

- Podríamos realizar alguna corrección.

```
error<-NA
acumulado<-NA
media.sin<-NA
media.corr<-NA
medias<-seq(0.5,4,length=20)

for(i in medias){
  proba<-dpois(0:20,i)
  acumu<-1-sum(proba[1:6])
  media<-sum((0:5)*proba[1:6])+6.0*(1-sum(proba[1:6]))
  media.sin<-c(media.sin,media)
  error<-c(error,i-media)
  acumulado<-c(acumulado,acumu)
}
acumulado<-acumulado[-1]
error<-error[-1]
media.sin<-media.sin[-1]

# Relación entre la media y el error
plot(medias,error,xlab=expression(lambda),ylab='Error')
title(main='Error que se comete con el truncamiento \n
en la estimación de la media')

# El problema es que hay que conocer la verdadera media
# pero si usamos el porcentaje de observaciones hasta el punto
# de truncamiento podemos aproximar la corrección.
plot(acumulado,error,xlab='Acumulado',ylab='Error')
```

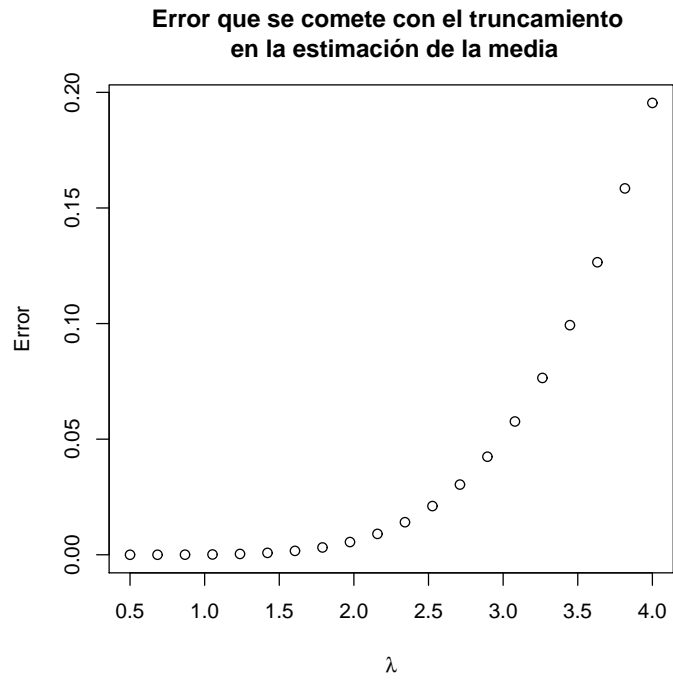


Figura 5.7. *Error absoluto cometido vs. media de la distribución de Poisson*

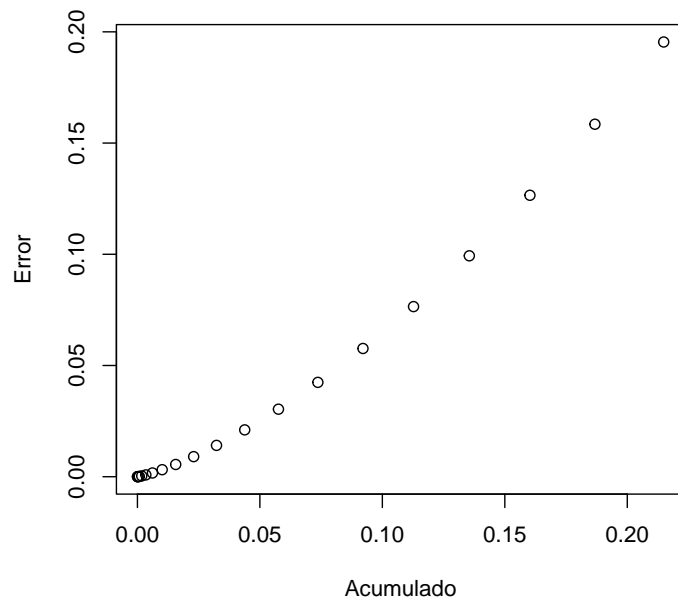


Figura 5.8. *Error cometido hasta el punto de truncamiento*

```
summary(modelo<-lm(error~acumulado+acumulado2))

## Call:
## lm(formula = error ~ acumulado + acumulado2)

## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.0010180 -0.0006270  0.0001269  0.0006803  0.0008440

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.0007817  0.0002460  -3.178  0.0055 **
## acumulado    0.4164908  0.0082370  50.563  <2e-16 ***
## acumulado2   2.3313949  0.0427119  54.584  <2e-16 ***
## ---

## Residual standard error: 0.0007206 on 17 degrees of freedom
## Multiple R-squared:  0.9999,    Adjusted R-squared:  0.9999
## F-statistic: 6.387e+04 on 2 and 17 DF,  p-value: < 2.2e-16
```

Podemos hallar la media haciendo la corrección mediante el ajuste cuadrático:

```
calcula.lambda<-function(proba){
  acumu<-1-sum(proba[1:6])
  media<-sum((0:5)*proba[1:6])+6.0*(1-sum(proba[1:6]))
  media.cor<-media-0.0007817+ 0.4164908*acumu
    + 2.3313949*acumu^2
  return(media.cor)
}

# Generación de la multinomial

temp<-scan()
170    250    300    180    60    35    5

res.multi<-rmultinom(2000,1000,temp)/1000
lambdas<-apply(res.multi,2,calcula.lambda)
hist(lambdas,freq=F,xlab=expression(lambda),
     main='Distribución a priori',ylab='Frecuencia')
summary(lambdas)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.679  1.808  1.835  1.836  1.863  1.979
```

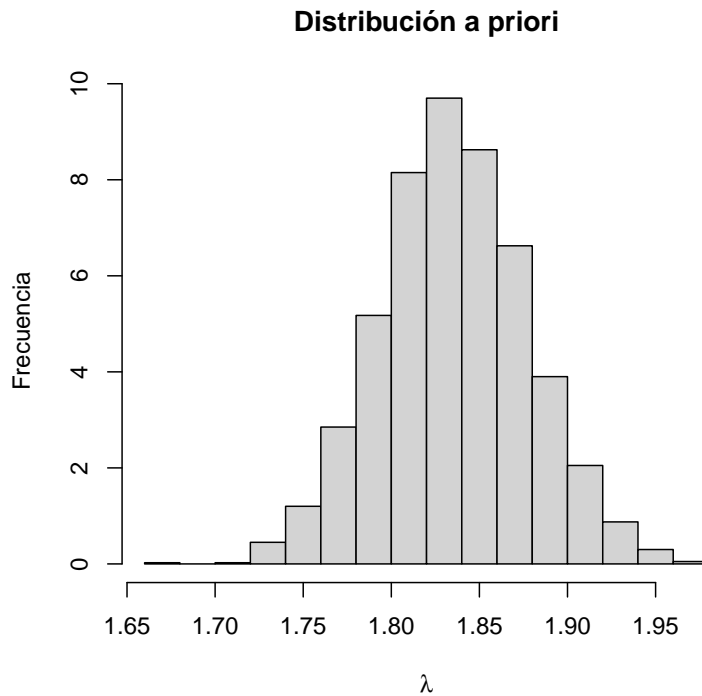


Figura 5.9. *Distribución a priori para la media de la Poisson.*

```
require(MASS)

fitdistr(lambdas, 'gamma')

##      shape      rate
## 1939.72904 1056.66458
## ( 58.85483) ( 32.06554)
```

Los parámetros de la gamma a priori serán:

$$\alpha = 1939.72904$$

$$\beta = 1056.66458$$

```
xx<-seq(1.5,2.2,length=100)
yy<-dgamma(xx,1939.72904,rate=1056.66458)
points(xx,yy,type='l',col='red')
```

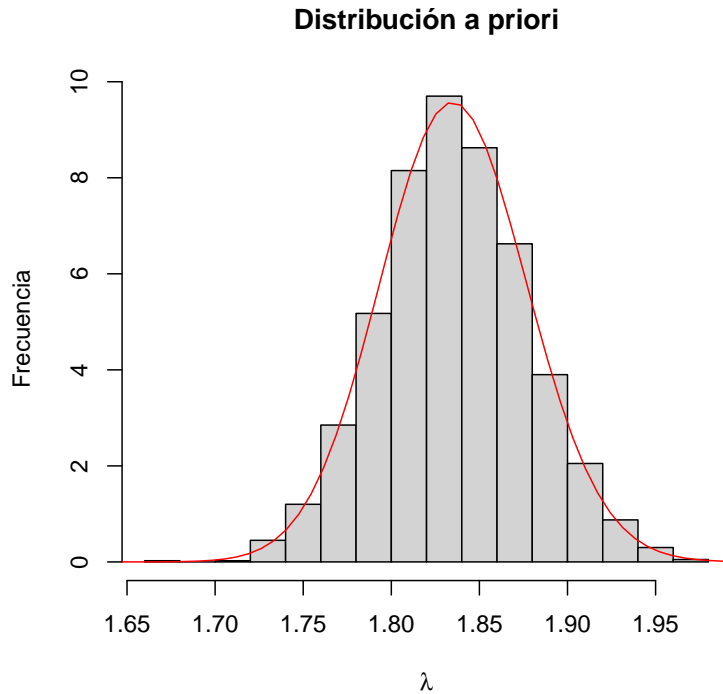


Figura 5.10. *Distribución a priori para la media de la poisson con modelo gamma ajustada.*

Especificación de los parámetros de una a priori normal para un parámetro de localización de una población normal

Preguntar con respecto a la localización es fácil: se puede conseguir información sobre la media o la mediana, o ambas, con el fin de validar estos valores. El problema es dar información sobre la precisión. Se pueden hacer preguntas directamente, pero el experto a veces confunde esto con valores poblacionales.

Una forma es realizar una elicitación de los parámetros poblacionales, la media (que es fácil) y sobre la dispersión poblacional.

- Determine el valor de la media, digamos μ_0 .
- Determine el máximo valor permisible para la población; esto puede hallarse asumiendo el percentil 99 % y denotemos este valor por $p_{0.99}$.
- Haga lo mismo con el mínimo; puede asumir el percentil 1 %, denotado por $p_{0.01}$.
- Bajo el supuesto de que la población es normal, rechequee los valores hallados en los dos pasos anteriores, ya que ellos exigen que $\mu_0 - p_{0.01} = p_{0.99} - \mu_0$. De aquí calculamos la varianza.

- A veces es conveniente generar *datos* de esta distribución poblacional y observar si corresponden a valores aceptables en la población que conocemos. Si, por ejemplo, aparecen valores mucho más extremos de lo que estamos dispuestos a aceptar, debemos revisar los valores elicitados; lo mismo en caso que los valores poblacionales simulados estén demasiado concentrados.
- Una vez determinados los valores anteriores, debemos comprobar la calidad de la información a priori que poseemos.
- Esto lo podemos hacer pensando en el tamaño muestral al que nuestro conocimiento equivale. Si estamos muy seguros del valor de la media que dimos, o sea μ_0 , escogemos un n_0 grande, digamos $n_0 = 1000$. En cambio, si no estamos seguros, seleccionamos valores pequeños. Nuestra a priori tendrá media μ_0 y precisión igual a $n_0/\text{varianza}$.

Tabla 5.2. *Tamaños de muestras correspondientes al nivel de seguridad del experto respecto a su información a priori*

n_0	Información a priori
0-10	Poca información a priori
11-20	Más bien poca
21-50	Buena información
50-100	Muy buen conocimiento
Más de 100	Excelente conocimiento.

Elicitación estatura

Suponga que queremos elicitar la estatura media de los hombres medellinenses.

- ¿Cuál es el valor más probable de la media? Respondo 170.
- ¿Cuál valor cree que divide la población en dos grupos tales que el 50% de los hombres está por debajo de esa estatura y el 50% por encima? Después de pensarlo un rato, diría que 170 me parece bien.
- Observe que los dos valores son iguales; esto sería bueno si el proceso de elicitación fuera mejor controlado.
- ¿Cuál sería la estatura tal que el 99% de los hombres está por debajo de ella? Pienso que si tengo una muestra de hombres de Medellín, el 99% estaría por debajo de 187.
- ¿Cuál sería el valor tal que el 99% estaría por encima de este? Creo que en una muestra de 100 hombres aceptaría hallar un hombre, a lo sumo, por debajo de 155.

- Ahora, $170 - 155 = 15 \neq 187 - 170 = 17$. Creo mucho en mi 187, así que ajusto 155 y acepto 153.
- Resuelvo

$$P\left(Z \leq \frac{187 - 170}{\sigma}\right) = 0.99 \implies \frac{187 - 170}{\sigma} = 2.326348 \quad (5.26)$$

$$\implies \sigma = 7.307592$$

```
x<-rnorm(100,mean=170,sd=7.307592)
head(x)
[1] 157.4084 172.7275 174.8296 170.3855 172.9300 161.4563
```

Si mi información a priori puedo compararla con un tamaño muestral de 20, entonces la a priori de μ será una normal con media 170 y precisión $20/7.307592^2$.

```
x <- seq(145, 200, by = 1)
y <- dnorm(x, mean = 170, sd = 7.307592)
y2 <- dnorm(x, mean = 170, sd = 7.307592 / sqrt(20))

plot(x, y, type = 'l', ylab = 'Densidad', xlab = 'Estatura', ylim = c(0, 0.3))
points(x, y2, type = 'l', lty = 2)

title(main = 'Distribución a priori del promedio')
legend(150, 0.30, legend = c('Distribución poblacional elicitada', 'Distribución a priori de la media'), lty=c(1, 2), cex=0.8, bty="n")
```

Distribución a priori del promedio

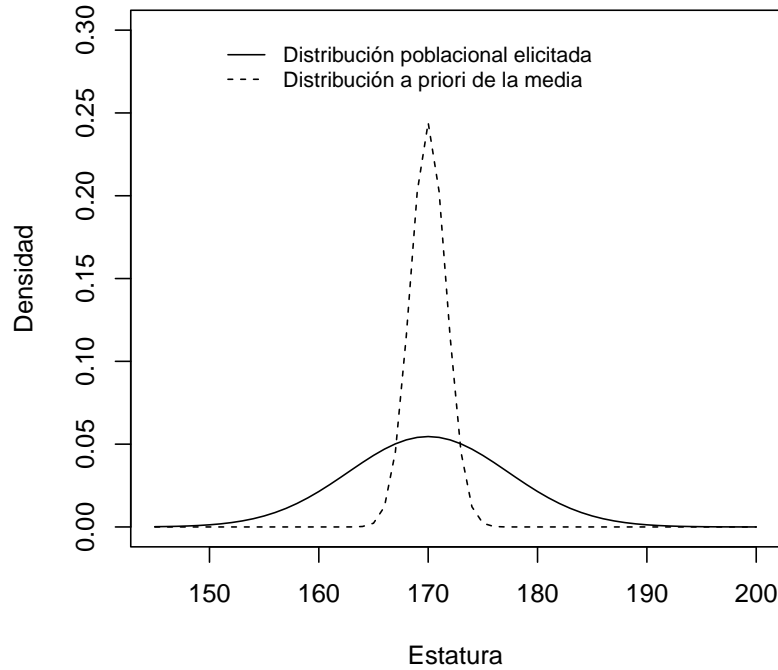


Figura 5.11. *Distribución a priori sobre la estatura promedio de los medellinenses obtenida usando una distribución poblacional elicitada y una información a priori equivalente a un tamaño muestral de 20*

Información muestral a priori equivalente (EPS)

Siendo parecido al anterior, en el caso de este procedimiento, realmente va asociado con alguno de los métodos anteriores (CDF y PDF paramétricos) y ayuda a calibrar los parámetros de la distribución escogida como a priori. El método EPS le pregunta al experto sobre una estimación del tamaño muestral en la que está basando su estimación inicial.

El método de la muestra a priori equivalente (EPS) y el método de la muestra futura hipotética (HFS) requieren una estimación directa de la proporción por parte del experto.

Mientras que los métodos EPS y HFS están restringidos generalmente a elicitar proporciones, los métodos CDF y PDF pueden usarse para elicitar la distribución de cualquier variable aleatoria continua. Los cuantiles, a menudo, se elicitan mediante la partición de subintervalos en otros de igual probabilidad. Este método de bisección es muy popular.

Hay situaciones de elicitación en las que el uso de esta técnica no tiene sentido, por ejemplo, en la elicitación de la edad de una persona.

Elicitación de los parámetros de un modelo lineal

Para ilustrar la adaptabilidad del método de muestras hipotéticas futuras, vamos a ilustrar la elicitación de los parámetros de un modelo lineal. Supongamos que estamos modelando el peso de hombres sanos, adultos y jóvenes (entre 20 y 25 años) de Medellín, explicando esto por medio de la estatura. O sea, el modelo a considerar es:

$$E(\text{Peso} | \text{Estatura}) = \beta_0 + \beta_1 \text{Estatura}$$

y los parámetros a los que le debemos construir la a priori son (β_0, β_1) .

Podemos escoger varios puntos en el espacio de diseño (la estatura) y a partir de estos construir las a priori predictivas de los pesos dados a estos valores. Para Medellín podríamos considerar hombres cuyas estaturas sean 165 cm, 170 cm y 175 cm. Para cada valor de estos necesitamos construir la distribución predictiva poblacional a priori.

Determinamos nuestro tamaño muestral equivalente a nuestro conocimiento sobre cada punto, digamos n_x , y extraemos $y_1^{(i)}(x), \dots, y_{n_x}^{(i)}(x)$. Teniendo esto ajustamos el modelo y obtenemos $\beta_0^{(i)}, \beta_1^{(i)}$, donde $i = 1, \dots, Nsim$. Luego de esta simulación tenemos $(\beta_0^{(1)}, \beta_1^{(1)}), (\beta_0^{(2)}, \beta_1^{(2)}), \dots, (\beta_0^{(Nsim)}, \beta_1^{(Nsim)})$. Usando esta información podemos determinar la distribución a priori estimada conjunta usando un kernel tipo histograma bivariable.

Los siguientes comandos demuestran cómo elicitar los parámetros del modelo lineal utilizando la distribución poblacional.

```
# Elicitación de los parámetros del modelo lineal vía la distribución
  n poblacional

x.s<-function(y.s,x1,x2,p1,p2){
  res<-lm(c(p1,p2)~c(x1,x2))$coefficients
  xx<-(y.s-res[1])/res[2]
  return(xx)
}
```

```

# Distribución a priori simulada a un nivel de x
resp.apriori.sim<-function(y.s,p.s,Nsim=1000,n.equ=10){
  simulados<-NULL
  y<-sort(runif(Nsim*n.equ))
  for(i in 2:length(y.s)){
    y.temp<-y[y<p.s[i]]
    n<-length(y.temp)
    y<-y[-(1:n)]
    temp<-x.s(y.temp,y.s[i-1],y.s[i],p.s[i-1],p.s[i])
    simulados<-c(simulados,temp)
  }# fin if
  simulados<-matrix(sample(simulados),ncol=n.equ)
  return(simulados)
}# fin resp.apriori.sim a un nivel de x

probas<-c(0,0.10,0.25,0.5,0.75,0.90,1)
# Para estatura=160
peso<-c(40,48,55,59,64,65,78)
X160<-resp.apriori.sim(peso,probas)

# Para estatura=170
peso<-c(55,63,65,68,73,75,90)
X170<-resp.apriori.sim(peso,probas)

# Para estatura=180
peso<-c(65,72,75,77,80,85,98)
X180<-resp.apriori.sim(peso,probas)

n.equ=10
YY<-cbind(X160,X170,X180)
estaturas<-c(rep(160,n.equ),rep(170,n.equ),rep(180,n.equ))

# función del modelo lineal
coefi.modelo<-function(y,x)lm.fit(cbind(1,x),y)$coefficients

Betas<-apply(YY,1,coefi.modelo,estaturas)

dim(Betas)
plot(Betas[1,],Betas[2,],
xlab=expression(beta[0]),
ylab=expression(beta[1]),main='Distribución Conjunta')

plot(density(Betas[1,]),main='Distribución a priori',
xlab=expression(beta[0]),ylab='Densidad')
plot(density(Betas[2,]),main='Distribución a priori',
xlab=expression(beta[1]),ylab='Densidad')

apply(Betas,1,mean)
apply(Betas,1,sd)
var(t(Betas))

```

```

require(ash)
x <- t(Betas)
ab <- matrix( c(-200,0,0,2), 2, 2) # interval [-5,5) x [-5,5)
nbin <- c( 20, 20) # 400 bins
bins <- bin2(x, ab, nbin) # bin counts,ab,nskip

m <- c(5,5)
f <- ash2(bins,m)
image(f$x,f$y,f$z)
contour(f$x,f$y,f$z,add=TRUE)

```

5.3. Apuestas

El uso de apuestas ha sido uno de los métodos más reconocidos para la elicitación de distribuciones subjetivas de parámetros con interpretación natural. Al experto se le presenta una situación como la siguiente:

- Apuesta 1: gana x pesos si el evento ocurre, pierde y pesos si el evento no ocurre.
- Apuesta 2: pierde x pesos si el evento ocurre, gana y pesos si el evento no ocurre.

La persona que realiza la elicitación cambia x y y hasta que la persona es indiferente ante la selección de cualquier apuesta.

Loterías hipotéticas. En este caso existen dos premios, uno muy atractivo y el otro no. El experto compara una lotería donde la probabilidad de ganar el premio atractivo es p (y ganar el no atractivo es $1 - p$) con una en la que se gana el premio atractivo si el evento de interés ocurre. El elicitor varía p hasta lograr que el experto sea indiferente a la selección entre las dos loterías. Pratt, Raiffa y Schlaifer (1964) presentan los supuestos inherentes a la técnica de las loterías. Una lotería se piensa en términos de una urna que contiene N bolas marcadas de 1 a N y que son indistinguibles en todos los aspectos, tal que la probabilidad de que una bola cualquiera salga es $1/N$. La idea básica es pensar en una lotería que tenga como resultado un evento A y otra lotería que tenga como resultado un valor en el intervalo $(0, \pi)$, $0 < \pi < 1$. Si la persona es indiferente ante las dos loterías (esto significa que para ella es indiferente quedarse con cualquiera de las dos), entonces la probabilidad de A es igual a π .

Winkler (1972) presenta el siguiente ejemplo sobre cómo funcionaría un proceso de elicitación usando loterías (la construcción formal desde un punto de vista constructivista de la definición de probabilidad subjetiva mediante el uso de loterías fue desarrollada en 1963 por Anscombe y Aumann):

Suponga que a usted debe escoger entre la lotería A y la lotería B.

Lotería A: Usted gana \$100 con probabilidad $\frac{1}{2}$.
Usted gana \$0 con probabilidad $\frac{1}{2}$.

Lotería B: Usted gana \$100 si llueve mañana.
Usted gana \$0 si no llueve mañana.

Se asume, dado que el premio es el mismo en ambas loterías, que usted preferiría la que le dé la mayor posibilidad de ganar el premio. Así, si usted escoge la lotería B, entonces debe sentir que la probabilidad de que llueva mañana es mayor que $\frac{1}{2}$; si usted escoge la lotería A, entonces sentiría que esta probabilidad es menor que $\frac{1}{2}$; si usted es indiferente entre las dos loterías, entonces siente que la probabilidad de que llueva mañana es igual a $\frac{1}{2}$. Ahora, considere las mismas loterías, excepto que las probabilidades en la lotería A se cambiaron a $\frac{1}{4}$ y $\frac{3}{4}$. Si usted aún prefiere la lotería A a la B, implica que siente que tiene una mayor oportunidad de ganar con A que con B, entonces su probabilidad subjetiva de lluvia es menor que $\frac{1}{4}$. Presumiblemente, usted puede estar cambiando las probabilidades en la lotería A hasta que sea indiferente a la selección entre la lotería A y la lotería B; si esto sucede cuando sus probabilidades sean 0.1 y 0.9, entonces su probabilidad subjetiva de lluvia es 0.1. De forma similar, usted puede determinar su probabilidad subjetiva de cualquier evento.

5.3.1. Un diálogo para la obtención de la proporción

El siguiente diálogo ilustra una posible forma de elicitar información sobre una proporción y es presentado en Raiffa (1970).

Analista: *Quisiera mostrarle una forma de obtener una distribución de probabilidad subjetiva acerca de una proporción π . Quiero elegir un contexto que sea lo suficientemente significativo para usted porque sus opiniones son las que interesan. Consideremos la población de médicos no abstemios en Colombia. Supongamos que π es la proporción de estos bebedores que consumieron más cerveza que aguardiente durante el último año. Por cierto, ¿sabe usted algo de los hábitos alcohólicos de los médicos?*

Cliente: No mucho. Conozco personalmente a tres o cuatro médicos, pero me imagino que los doctores no serán muy diferentes de los abogados, los dentistas o los ingenieros. El problema es que no sabría responder a su pregunta para cualquiera de esos grupos. No tengo ni la más remota idea del valor de π .

Analista: *Bueno. Esto es justamente lo que quería.*

Cliente: Supongo que quiere que haga una predicción óptima de π . No sé si podría hacerlo.

Analista: *No, no es eso lo que quiero. De hecho, no creo que tenga sentido hablar de una predicción “óptima”. ¿Óptima para qué? Déjeme hacerle unas preguntas de precalentamiento. ¿Cree usted que es probable que π sea menor o mayor que 0.10?*

Cliente: Mayor, seguro.

Analista: *¿Es probable que π sea mayor o menor que 0.90?*

Cliente: Menor.

Analista: *Estas preguntas eran fáciles. Vea usted, ya sabe algo sobre π . Ahora quiero que piense bien antes de contestarme. Deme una proporción tal que sea extremadamente difícil decidir si es probable que π sea mayor o menor que ese valor. En otras palabras, quiero que me dé un valor para el que crea que es igualmente probable que π sea mayor o menor que él.*

Cliente: (Después de pensar un rato). Yo diría que 0.60. Pero no estoy nada seguro. Creo que la mayoría de los médicos prefiere tomar cerveza.

Analista: *No se preocupe demasiado; si quiere cambiar de opinión más tarde, no hay inconveniente. Me acaba de decir que cree que es igualmente probable que π sea mayor o menor que 0.60.*

Cliente: Exactamente. Pero no me pida que defina lo que significa “igualmente probable”.

Analista: *Por “igualmente probable”, en este contexto, quiero decir que usted es indiferente entre recibir un premio muy deseable condicionado a que π sea menor que 0.60, y recibir este mismo premio condicionado a que π sea mayor que 0.60. O, de forma más dramática, si su vida dependiera de ello, optaría igualmente por un $\pi \leq 0.60$ que por un $\pi \geq 0.60$. ¿Está siguiéndome?*

Cliente: Hasta ahora sí.

Analista: *Esencialmente, me acaba de decir, y usted lo cree, que 0.60 divide el intervalo de cero a uno en dos partes igualmente probables, en su opinión. Ahora voy a pedirle que repita este proceso de dividir imaginariamente diferentes intervalos en dos partes igualmente probables. Por ejemplo, ¿qué cree usted que es más probable, que π sea menor que 0.20 o que esté comprendido entre 0.20 y 0.60?*

Cliente: Entre 0.20 y 0.60.

Analista: *¿Entre cero y 0.58, o entre 0.58 y 0.60?*

Cliente: Entre 0 y 0.58.

Analista: De acuerdo. Ahora deme un número para el que crea que es igualmente probable que π esté comprendido entre cero y ese número y ese número y 0.60.

Cliente: ¿Qué ocurre si π es mayor que 0.60?

Analista: Tal como están las cosas usted pierde. Mire, si me dice que el número es π^* , esto quiere decir que sus probabilidades de ganar el premio son igualmente buenas si elige el intervalo de cero a π^* como lo son si elige el intervalo de π^* a 0.60. Si π es mayor que 0.60, no obtendría el premio sea cual sea el lado de π^* que elija, porque π no estaría en esos intervalos.

Cliente: De acuerdo. Veamos... Diré que 0.50 divide el intervalo de cero a 0.60 en dos partes igualmente probables.

Analista: ¿Le hubiera sido más fácil si, una vez que me había dado el número 0.60 le hubiera hecho la última pregunta de esta otra manera?: Mire, suponga que le digo que π es menor que 0.60. ¿Cómo dividiría ahora el intervalo de cero a 0.60 en dos partes igualmente probables?

Cliente: ¿Son iguales las preguntas?

Analista: Creo que sí. Piénselo.

Cliente: Supongo que son iguales. La segunda me parece más fácil, pero las cosas siempre me parecen más fáciles la segunda vez.

Analista: Continuemos. Imagínese que le digo que π es mayor que 0.60. ¿Cómo dividiría el intervalo 0.60 a 1 en dos partes igualmente probables?

Cliente: 0.70. El intervalo de 0.60 a 0.70 me parece igualmente probable que los valores superiores a 0.70. Pero realmente me siento un poco incómodo acerca de 0.50 y 0.70 porque el 0.60 es muy poco firme. Me parece como si estuviera construyendo sobre una esponja. Espero que se dé cuenta de que estos números son muy poco firmes.

Analista: Me doy cuenta de ello. ¡Animo! Ahora me ha dado tres números, 0.60, 0.50 y 0.70. Voy a dibujar un intervalo de cero a 1 y a colocar estos puntos en él.

0  1

Me acaba de decir que, en lo que a usted respecta, cree que es igualmente probable que π esté situado en cualquiera de los cuatro intervalos [0 a 0.50], [0.50 a 0.60], [0.60 a 0.70] y (0.70 a 1.00].

Cliente: Sí, creo que eso es lo que he dicho.

Analista: *Ahora voy a comprobarlo. No pretendo confundirle ni molestarle, pero es importante considerar estas cuestiones desde todos los ángulos. Por ejemplo, ¿preferiría usted apostar que π está en el intervalo [0.50 a 0.70] o apostar que está fuera de este intervalo?*

Cliente: Creo que apostaría que está dentro del intervalo. Pero estoy siendo inconsistente, ¿no?

Analista: *Sí, lo está siendo, pero es normal. Quiero que piense más sobre esto. Es una ventaja que trate conscientemente de ser consistente.*

Cliente: Bien, no quiero cambiar el 0.60. Me siento más inseguro con el 0.70. Supongo que estaría dispuesto a admitir un 0.68. En lo que a mí respecta, hay una probabilidad del 50-50 de que π esté en el intervalo [0.50 a 0.68].

Analista: *¿Estaría dispuesto a decir que es igualmente probable que π esté en el intervalo [0.60 a 0.68] o en el intervalo [0.68 a 1]?*

Cliente: De acuerdo, lo admitiré. Pero, si lo hiciéramos todo otra vez y si borrara esta conversación de mi memoria, puedo imaginar que, en lugar de terminar con los números 0.50, 0.60 y 0.68, podría haber terminado con números como 0.52, 0.64 y 0.74.

Analista: *Bien, ¿podría usted imaginar terminar con números tales como 0.20, 0.40 y 0.55?*

Cliente: No. Realmente no. Pero, ¿qué haría si yo dijera sí?

Analista: *Seguiría presionándole y utilizaría algún sistema de promedios que separara aún más los tres que me ha dado. Pero, continuemos. Me referiré al número 0.60 como su percentil 0.50, al número 0.50 como su percentil 0.25 y al número 0.68 como su percentil 0.75. Necesito unos cuantos porcentajes más. ¿Cómo dividiría usted el intervalo [0 a 0.50] en dos partes igualmente probables?*

Cliente: 0.42.

Analista: *Ahora divida el intervalo [0 a 0.42]*

Cliente: Me está forzando demasiado.

Analista: *Bien. Suponga que yo le dijera que π es menor que 0.42. ¿Preferiría apostar por el intervalo [0 a 0.21] o por el intervalo [0.21 a 0.42]?*

Cliente: Por el último, por supuesto. De acuerdo, utilice 0.36.

Analista: *Ahora pasemos al extremo superior. Divida [0.68 a 1.00].*

Cliente: Utilice 0.75.

Analista: De acuerdo. Divida [0.75 a 1.00].

Cliente: Utilice 0.80.

Analista: Resumamos sus respuestas en una tabla (véase la tabla 5.3).

Tabla 5.3. *Distribución subjetiva del cliente*

Percentil	Valor del percentil
0.0625	0.36
0.125	0.42
0.25	0.50
0.50	0.60
0.75	0.68
0.875	0.75
0.9375	0.80

Esta discusión puede continuar por largo tiempo y, a medida que el cliente es más y más consciente del proceso de elicitación, su información es cada vez mejor.

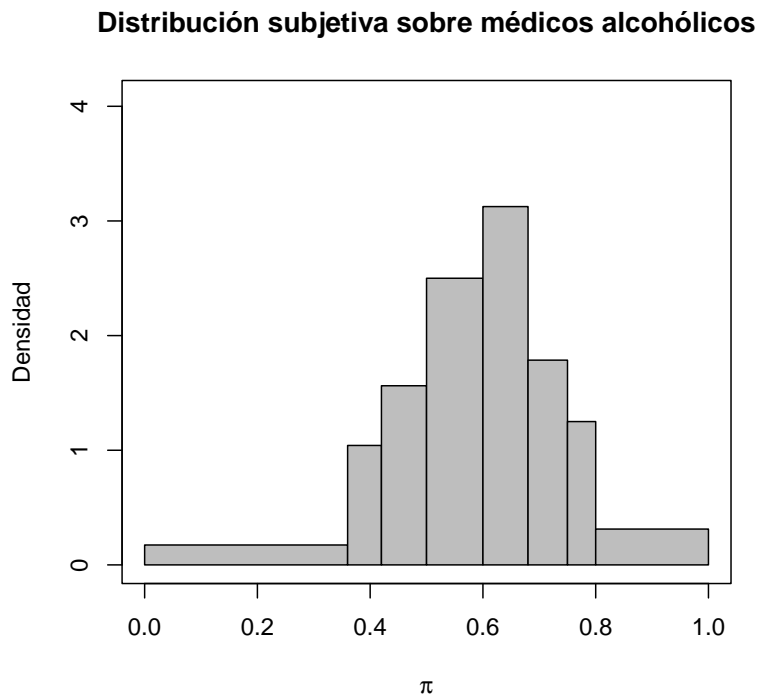


Figura 5.12. *Distribución a priori obtenida del diálogo llevado a cabo con relación al porcentaje de los médicos alcohólicos*

```

temp<-scan()
0.0625  0.36
0.125  0.42
0.25   0.50
0.50   0.60
0.75   0.68
0.875  0.75
0.9375 0.80

temp<-matrix(temp,ncol=2,byrow=T)
temp<-rbind(c(0,0),temp,c(1,1))
Area<-temp[2:9,1]-temp[1:8,1]
base<-temp[2:9,2]-temp[1:8,2]
Altura<-Area/base

plot(1,1,ylim=c(0,max(Altura)*1.3),xlim=c(0,1),type='n',
     ylab='',xlab='')
title(ylab='Densidad',xlab=substitute(pi))
title(main='Distribución Subjetiva sobre Médicos Alcohólicos')
for(i in 1:length(Altura)){
  rect(temp[i,2],0,temp[i+1,2],Altura[i],col='grey')
}

```

5.3.2. Problemas con el uso de loterías

Las primeras aproximaciones al problema de elicitar probabilidades de eventos surgieron en escuelas de Administración de Negocios en Estados Unidos. Obviamente es un ambiente propicio a la utilización del concepto de función de utilidad personal, algo que surge en forma natural en el área de toma de decisiones. El problema es que la función de utilidad personal debería determinarse antes de poder elicitar las probabilidades personales. Ante esta situación, se toman alternativas tales como asumir que la función de utilidad es lineal, algo que en muchas circunstancias no es cierto.

La asignación de probabilidades subjetivas vía apuestas sufre de diferentes problemas si son utilizadas con sujetos ingenuos. El esfuerzo cognitivo de la estimación de esperanzas es mayor que la especificación directa de probabilidades. El tamaño de la apuesta puede afectar el uso de la técnica. La deseabilidad del premio puede sesgar las respuestas del sujeto. Estructurar las preguntas puede ser difícil si se necesitan muchas asignaciones. El método de la lotería no sufre el problema de la utilidad en el caso de las apuestas; sin embargo, es un procedimiento tedioso si se necesita responder un gran número de preguntas (Chesley, 1975).

5.4. Elicitación en el modelo lineal

Garthwaite (2007) y Garthwaite y Dickey (1988) proponen un método de elicitación en el modelo lineal. Otros métodos apelan a una configuración más intuitiva, como crear una imagen mental de la población. Esto permite a los expertos obtener muestras mentales de esta población que, con suerte, tendrá una muestra que preserve la estructura de la población. La calidad de su información personal se cuantifica en términos del tamaño de muestra equivalente. Cristov et al. (2017) aplican varios métodos de elicitación, pero este estudio no permite producir conclusiones generales debido a sus limitaciones extremas. Renooij y Wittteman (1999) y Wittteman y Renooij (2003) proponen el uso de la expresión verbal para cuantificar las probabilidades personales.

La elicitación del vector de parámetros de un modelo lineal normal, cuando se desconoce, es importante. Kadane et al. (1980) presentan un método heurístico para elicitación de los hiperparámetros en el modelo lineal normal en el caso de una a priori conjugada. Varios autores desarrollaron un programa de computador para ayudar con esta tarea (James et al., 2010). Un problema con esta aproximación y el programa es que requiere la obtención directa de los parámetros del modelo. Esto podría ser fácil si hubiera estudios previos que proporcionen información directa sobre ellos.

Nuestra propuesta (véase Correa-Morales y Barrera-Causil 2021) se basa en la construcción de una imagen mental de la población de respuestas en varios puntos específicos de las variables explicativas en el modelo, y se obtiene la media y la varianza en cada uno de estos puntos. Esta tarea es mucho más fácil que la obtención directa de los parámetros del modelo. El algoritmo se describe a continuación:

5.4.1. El algoritmo

Asuma que estamos interesados en el siguiente modelo lineal (esta propuesta se puede extender a modelos más complejos).

$$y \sim \beta_0 + \beta_1 \cdot x + \epsilon, \quad (5.27)$$

donde ϵ se asume que sigue una distribución normal con media 0 y precisión $\tau = \frac{1}{\sigma^2}$. Algunos estadísticos bayesianos prefieren trabajar con la inversa de la varianza y lo llaman la *precisión* (DeGroot, 1970). Esto facilita el trabajo de álgebra para distribuciones conjugadas y evita trabajar con la distribución gamma inversa en lugar de la distribución gamma tradicional. Nuestro problema básico es la determinación de la distribución conjunta previa de los parámetros del modelo, digamos $\xi(\beta_0, \beta_1, \sigma)$.

Esta propuesta está dividida en dos pasos. En el primero, algunos cuantiles de la distribución de la variable respuesta en el modelo de regresión son elicitados con respecto a un valor particular de la variable explicativa. En este paso se obtienen

también la media, varianza y el correspondiente tamaño de muestra equivalente. Note que este proceso de elicitación de los cuantiles de la variable aleatoria es más fácil de llevar a cabo que la elicitación de los parámetros del modelo directamente (Demuyck, 2013).

En el segundo paso, se realiza un proceso de simulación para generar muestras de la variable respuesta considerando sus correspondientes puntos de diseño, esto con el objetivo de estimar, de manera repetida, los parámetros del modelo. Aquí, es de notar que es más fácil para los expertos dar opiniones alrededor de un valor central, que alrededor de cualquier otro en el rango de posibles valores de una variable. Este argumento se fundamenta en el hecho de que es más fácil describir eventos que ocurren con mucha frecuencia, que aquellos que son raros. De esta manera, en esta propuesta de Correa-Morales y Barrera-Causil (2021), se recomienda establecer como puntos de diseño aquellos valores cercanos a la media de la variable explicativa.

Paso 1. Elicitación

Primero, pedimos al experto que elija un punto de la variable explicativa en su rango lógico, digamos x_0 . Luego, se le pide que imagine la distribución de las respuestas en este nivel de la variable explicativa x_0 y elija un valor de la respuesta en el rango lógico de esta, digamos y_0 .

Posteriormente, le presentamos al experto la siguiente situación: si tenemos una muestra representativa de esta población condicional de tamaño n_0 , ¿cuántas observaciones considerará que están por debajo de y_0 ? Llame a este número n_{y_0} . Luego, pregúntele cuál sería el número mínimo de observaciones que acepte por debajo de y_0 en la muestra. Llame a este número k_l . Ahora, pregunte al experto por la cantidad máxima de observaciones que acepte por encima de y_0 en la muestra. Llame a este número k_u .

En cada punto obtenido, el experto debe producir un tamaño de muestra hipotético que refleje su confianza en esta distribución provocada. Al final del proceso vamos a tener:

$$\{\xi(y_i | x_{i1}, x_{i2}, \dots, x_{ip}), n_i\}_{i=1}^k \tag{5.28}$$

Supongamos ahora que obtenemos los siguientes valores.

Pregunta 1:

Nivel de la variable explicativa ($X = x_0$)			
	Muestra hipotética	Número dado	Percentil esperado
Lo más posible	n_0	n_{y_0}	$E[\text{Percentil}] = \frac{n_{y_0}}{n_0}$
Más posible	n_0	$n_{y_0} + k_u$	$E_u[\text{Percentil}] = \frac{n_{y_0} + k_u}{n_0}$
Lo menos posible	n_0	$n_{y_0} - k_l$	$E_l[\text{Percentil}] = \frac{n_{y_0} - k_l}{n_0}$

Pregunta 2:

Nivel de la variable explicativa ($X = x_0$)			
	Muestra hipotética	Número dado	Percentil esperado
Lo más posible	n_1	n_{y_1}	$E[\text{Percentil}] = \frac{n_{y_1}}{n_1}$
Más posible	n_1	$n_{y_1} + k_u$	$E_u[\text{Percentil}] = \frac{n_{y_1} + k_u}{n_1}$
Lo menos posible	n_1	$n_{y_1} - k_l$	$E_l[\text{Percentil}] = \frac{n_{y_1} - k_l}{n_1}$

Ahora, asuma que para el mismo valor de la variable explicativa x_0 , y bajo n_0 , podemos estandarizar y_0 , tal que:

$$\frac{y_0 - \mu_{x_0}}{\sigma} \approx \Phi^{-1} \left(\frac{n_{y_0}}{n_0} \right). \quad (5.29)$$

De forma similar, pero bajo n_1 , tenemos:

$$\frac{y_1 - \mu_{x_0}}{\sigma} \approx \Phi^{-1} \left(\frac{n_{y_1}}{n_1} \right). \quad (5.30)$$

De aquí, se resuelven las siguientes ecuaciones para μ_{x_0} y σ :

$$\sigma = \frac{y_0 - y_1}{\Phi^{-1} \left(\frac{n_{y_0}}{n_0} \right) - \Phi^{-1} \left(\frac{n_{y_1}}{n_1} \right)} \quad (5.31)$$

$$\mu_{x_0} = y_0 - \left(\frac{y_0 - y_1}{\Phi^{-1} \left(\frac{n_{y_0}}{n_0} \right) - \Phi^{-1} \left(\frac{n_{y_1}}{n_1} \right)} \right) \Phi^{-1} \left(\frac{n_{y_0}}{n_0} \right). \quad (5.32)$$

Seguidamente, se halla el tamaño de muestra equivalente (n_{equ}) que refleje el conocimiento del experto respecto a la información proporcionada. Para esto, se utiliza un intervalo de confianza binomial para una proporción del tipo:

$$\hat{\pi} \pm 1.96 \sqrt{\frac{\hat{\pi} (1 - \hat{\pi})}{n_{equ}}}. \quad (5.33)$$

Si asumimos que el límite superior es $\frac{n_{y_0} + k_u}{n_0}$, entonces:

$$\frac{n_{y_0} + k_u}{n_0} = \frac{n_{y_0}}{n_0} + 1.96 \sqrt{\frac{\frac{n_{y_0}}{n_0} \left(1 - \frac{n_{y_0}}{n_0} \right)}{n_{equ}}}. \quad (5.34)$$

Así, el (n_{equ}) equivalente de la creencia del experto en este punto de diseño se aproxima como sigue:

$$n_{equ1} = \left(\frac{1.96}{k_u} \right)^2 n_{y_0} (n_0 - n_{y_0}). \quad (5.35)$$

Si asumimos que el límite inferior es $\frac{n_{y_0}-k_l}{n_0}$, entonces:

$$\frac{n_{y_0} - k_l}{n_0} = \frac{n_{y_0}}{n_0} - 1.96 \sqrt{\frac{\frac{n_{y_0}}{n_0} \left(1 - \frac{n_{y_0}}{n_0}\right)}{n_{equ}}}. \quad (5.36)$$

El (n_{equ}) equivalente del experto en este punto de diseño es aproximado de la siguiente manera:

$$n_{equ2} = \left(\frac{1.96}{k_l}\right)^2 n_{y_0} (n_0 - n_{y_0}). \quad (5.37)$$

Dado que preferimos ser cautelosos con la información proporcionada por el experto, seleccionamos el siguiente tamaño de muestra equivalente: $n_{x_0} = \min\{n_{equ1}, n_{equ2}\}$.

Asumiendo x_1 como un punto de diseño diferente para la variable explicativa, se sigue un proceso similar al descrito.

En resumen, para dos puntos de la variable explicativa, digamos x_0 y x_1 , tenemos dos tablas como las presentadas. De las dos filas *Lo más posible*, se obtienen las siguientes ecuaciones:

$$\begin{aligned} \frac{x_0 - \mu}{\sigma} &= \Phi^{-1}\left(\frac{n_{x_0}}{n_0}\right), \text{ y} \\ \frac{x_1 - \mu}{\sigma} &= \Phi^{-1}\left(\frac{n_{x_1}}{n_1}\right). \end{aligned} \quad (5.38)$$

5.4.2. Ilustración

Correa-Morales y Barrera-Causil (2021) ilustran este proceso de elicitación asumiendo que se tiene una población de jugadores de fútbol y el propósito es modelar el peso (y_0) promedio de estos usando sus estaturas (x_0) como la variable explicativa. De esta manera, si se tiene un tamaño de muestra n_0 de jugadores con $x_0 = 170$ cm, se puede preguntar al experto lo siguiente: como máximo, ¿cuántos jugadores cree usted pesan menos de 70 kg? (su respuesta será n_{y_0}). ¿Cuál es el número máximo de jugadores de fútbol que usted aceptaría en esta muestra particular que pesen 70 kg? (Su respuesta será k_u). De forma similar se obtiene el mínimo y se denota por k_l .

Ahora bien, si se asume otra muestra hipotética, digamos n_1 , las preguntas pueden ser como sigue: en una muestra de jugadores de fútbol que miden 170 cm, como máximo, ¿cuántos de ellos pesan menos de 80 kg?, ¿cuál considera usted que es el número máximo de jugadores de fútbol de esta muestra que pesen 80 kg?, ¿cuál es el mínimo que aceptaría? Este proceso se puede repetir el número de veces que el elicitador considere.

Una vez se tienen las poblaciones para diferentes puntos de las variables explicativas con sus respectivos tamaños de muestra equivalente, se construye la distribución a priori conjugada de los parámetros de regresión.

Paso 2. Simulación

1. Para la i -ésima población especificada, obtenga un tamaño de muestra simulado n_{x_i} , digamos \mathbf{y}_i .
2. Construya:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad (5.39)$$

y

$$\mathbf{X} = \begin{bmatrix} \mathbf{1} & \mathbf{x}_1 \\ \mathbf{1} & \mathbf{x}_2 \end{bmatrix}. \quad (5.40)$$

3. Estime el modelo $y = \beta_0 + \beta_1 x + e$, donde $E[e] = 0$ y $Var(e) = \sigma^2$. Mantenga $(\hat{\beta}_0, \hat{\beta}_1, \hat{\sigma}^2)$.
4. Repita muchas veces los pasos de arriba, digamos M .

Una vez que hemos hecho el proceso anterior, estimamos los parámetros de una distribución multinormal para los β y una gamma para la varianza. La distribución a priori conjugada es la distribución normal-inversa-gamma, la cual es expresada como sigue:

$$\begin{aligned} \xi(\beta, \sigma^2) &= \xi(\beta | \sigma^2) \xi(\sigma^2) = N(\mu_\beta, \sigma^2 \Sigma_\beta) \times IG(\alpha_0, \gamma_0) \\ &\propto \left(\frac{1}{\sigma^2}\right)^{(\alpha_0 + p/2 + 1)} \times \exp \left[-\frac{1}{\sigma^2} \left\{ \gamma_0 + \frac{1}{2} (\beta - \mu_\beta)' \Sigma_\beta (\beta - \mu_\beta) \right\} \right], \end{aligned} \quad (5.41)$$

donde p es el número de regresores y los hiperparámetros α_0 y γ_0 deben ser estimados usando una distribución a priori no informativa o utilizando elicitación.

La determinación de α_0 y β_0 en muchas situaciones está rodeada de incertidumbre. En este caso, los investigadores establecen una distribución en $(\alpha_0, \gamma_0)'$, que se denomina distribución hiper a priori.

Finalmente, con este método de elicitación propuesto, se lleva a cabo la inferencia bayesiana considerando a priori informativas.

5.5. Ejemplo hipotético

Para ilustrar este procedimiento, asumamos que estamos interesados en modelar el peso de estudiantes universitarios basados en su estatura y el sexo, usando un modelo lineal.

$$\text{Peso} = \beta_0 + \beta_1 \text{Sexo} + \beta_2 \text{Estatura} + e, \quad (5.42)$$

donde $e \sim N(0, \sigma^2)$. Si seleccionamos cuatro puntos para elicitación del peso de los estudiantes (dos puntos para cada género, digamos $x_{1M} = 170$ cm, $x_{2M} = 180$ cm, $x_{1F} = 160$ cm y $x_{2F} = 170$ cm), elicitamos los parámetros de cuatro poblaciones de distribución normal; aquí, la elicitación de la media y la varianza de cada distribución puede ser más fácil que la elicitación de los parámetros del modelo.

Usando cada punto de diseño y diferentes muestras hipotéticas, se puede preguntar al experto como sigue:

Para estudiantes de sexo masculino. Con una muestra de tamaño n_0 de estudiantes universitarios de 170 cm de estatura, ¿como máximo, cuántos estudiantes de esta muestra cree usted que pesan menos de 70 kg?, ¿cuál considera que es el máximo número de estudiantes en esta muestra de hombres que pesen 70 kg?, ¿cuál es el mínimo que usted considera para este caso?

Ahora, usando las dos tablas presentadas un poco más atrás, de las dos filas *Lo más posible*, podemos obtener la media y la varianza de cada población, y generar los correspondientes y_{ki} , $k = 1, 2, 3, 4$, y $i = 1, 2, \dots, m$ de $N_{\{M, x=170\}}(\mu_1, \sigma_1^2)$, $N_{\{M, x=180\}}(\mu_2, \sigma_2^2)$, $N_{\{F, x=160\}}(\mu_3, \sigma_3^2)$, y $N_{\{F, x=170\}}(\mu_4, \sigma_4^2)$.

Ahora, con:

$$\mathbf{Y} = \begin{bmatrix} y_{11} \\ \vdots \\ y_{n_{x_0}1} \\ y_{12} \\ \vdots \\ y_{n_{x_1}2} \\ y_{13} \\ \vdots \\ y_{n_{x_2}3} \\ y_{14} \\ \vdots \\ y_{n_{x_3}4} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & 170 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 170 \\ 1 & 1 & 180 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 180 \\ 1 & 0 & 160 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 160 \\ 1 & 0 & 170 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 170 \end{bmatrix}. \quad (5.43)$$

Luego, siguiendo el tercer paso del algoritmo, estimamos $(\beta_0, \beta_1, \beta_2, \sigma^2)$. Finalmente, con los pasos 2, 3, y 4, obtenemos la distribución a priori conjugada. En el siguiente enlace se encuentra el código para esta ilustración que presentamos en

Correa-Morales y Barrera-Causil (2021).

(<https://github.com/cbarrera2101/Code-elicitation-of-the-parameters-for-LM.git>),

Elicitación de la distribución a priori de los parámetros del modelo lineal simple

```
# Determinación de los parámetros de una normal a partir de de una
  muestra de tamaño n representada en intervalos

estimaNormal <- function(parametros,frecu=frecu,limites=limites){
  media<-parametros[1]
  vari<-parametros[2]
  dev.tip<-sqrt(vari)
  proba.esti<-frecu/sum(frecu)
  error<-0.0
  for(i in 1:length(frecu)){
    error<-error+(pnorm(limites[i+1],mean=media,sd=dev.tip)-
  pnorm(limites[i],mean=media,sd=dev.tip)-proba.esti[i])^2
  }
  return(error)
}
```

```
# Determinación de la a priori de beta0, beta1 y sigma2 en el modelo
  Peso=beta0+beta1 Estatura + error.
# Consideremos los hombres que miden exactamente 170cm. De 1000
  hombres en una muestra hipotética
limites<-c(45,55,60,70,80,90)
frecus<-c(0,10,450,440,0)

(resu.170<-optim(c(70,25),estimaNormal,method='L-BFGS-B',
lower=c(40,0.01),upper=c(100,100),
limites=limites,frecu=frecus))

param.opt170<-resu.170$par
```

```
## $par
## [1] 69.94044 15.57328

## $value
## [1] 5.738812e-05

## $counts
## function gradient
##      13      13

## $convergence
## [1] 0

## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"
```

```

# Consideremos los hombres que miden exactamente 180cm. De 1000
  hombres en una muestra hipotética
limites<-c(55,65,70,75,80,85,95,110)
frecus<-c(0,10,100,400,450,35,5)

(resu.180<-optim(c(80,25),estimaNormal,method='L-BFGS-B',
lower=c(50,0.01),upper=c(110,100),
limites=limites,frecu=frecus))

param.opt180<-resu.180$par

```

```

## $par
## [1] 80.08554 11.48700

## $value
## [1] 0.003506067

## $counts
## function gradient
##      13      13

## $convergence
## [1] 0

## $message
## [1] "CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH"

```

Función que simula una muestra de Y a dos niveles de la covariable y estima los parámetros del modelo

```

calcula.parame.model<-function(ii,nivel1,nivel2,
media1,media2,var1,var2,
n.seg1,n.seg2){
  xx<-c(rep(nivel1,n.seg1),rep(nivel2,n.seg2))
  yy<-c(rnorm(n.seg1,mean=media1,sd=sqrt(var1)),
        rnorm(n.seg2,mean=media2,sd=sqrt(var2)))
  res.lm<-lm(yy~xx)
  res<-c(res.lm$coefficients,mean((res.lm$residuals)^2))
  return(res)
}

```

Niveles de seguridad del experto a cada nivel de X

```
# Nivel de seguridad en nuestro conocimiento n.seg170=25
n.seg170<-15

# Nivel de seguridad en nuestro conocimiento n.seg180=20
n.seg180<-10
```

Definición del número de simulaciones y la simulación usando la función sapply()

```
Nsim<-1000000

resu<-matrix(sapply(rep(00,Nsim),calcula.parame.modelo,
170,180,param.opt170[1],param.opt180[1],
param.opt170[2],param.opt180[2],
n.seg170,n.seg180),ncol=3,byrow=T)
```

Medidas descriptivas

```
dim(resu)
## [1] 1000000      3

cov(resu)
##           [,1]      [,2]      [,3]
## [1,] 668.0207367 -3.819631848  0.204675186
## [2,] -3.8196318  0.021857879 -0.001174016
## [3,]  0.2046752 -0.001174016 14.687252337

cor(resu)
##           [,1]      [,2]      [,3]
## [1,] 1.000000000 -0.999591430  0.002066331
## [2,] -0.999591430  1.000000000 -0.002072048
## [3,]  0.002066331 -0.002072048  1.000000000

colMeans(resu)
## [1] -102.533666  1.014553  12.856778
apply(resu,2,median)
## [1] -102.565012  1.014658  12.475008
```

Gráficos de las distribuciones a priori marginales simuladas

```
hist(resu[,1],main=expression(beta[0]),ylab='Frecuencia  
relativa',xlab=expression(beta[0]),  
freq=F,col='gray')
```

```
hist(resu[,2],main=expression(beta[1]),ylab='Frecuencia  
relativa',xlab=expression(beta[1]),  
freq=F,col='gray')
```

```
hist(resu[,3],main=expression(sigma^2),ylab='Frecuencia  
relativa',xlab=expression(sigma^2),  
freq=F,col='gray')
```

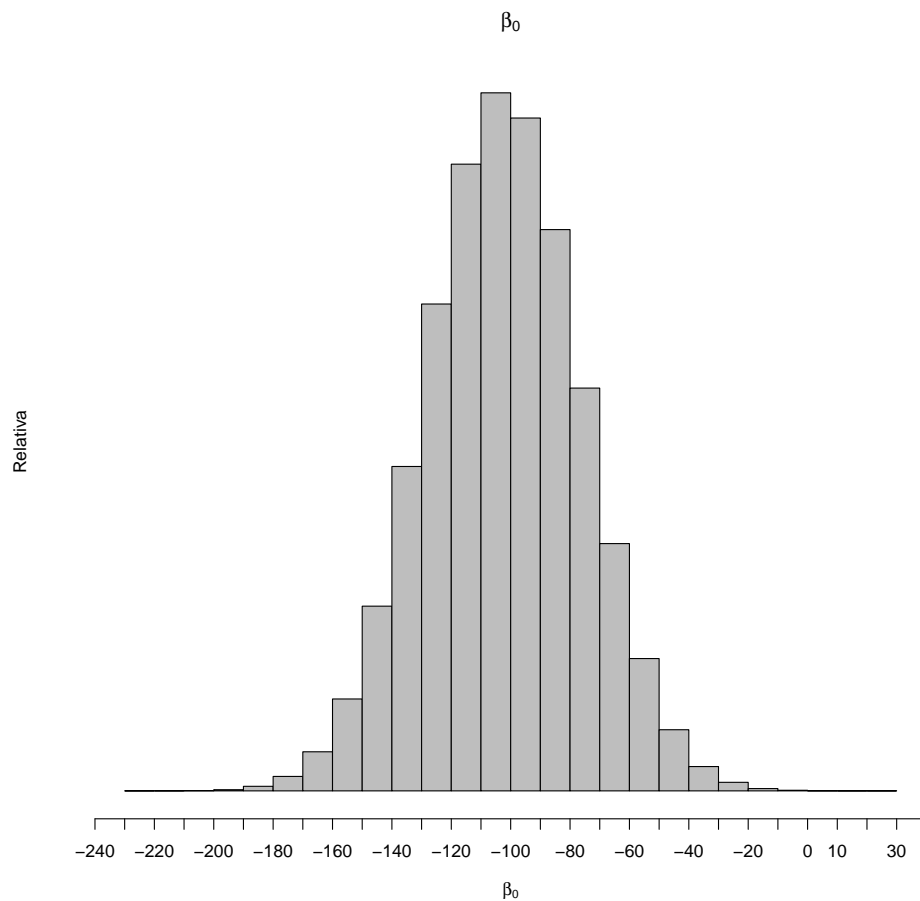


Figura 5.13. *Distribución a priori marginal simulada para β_0*

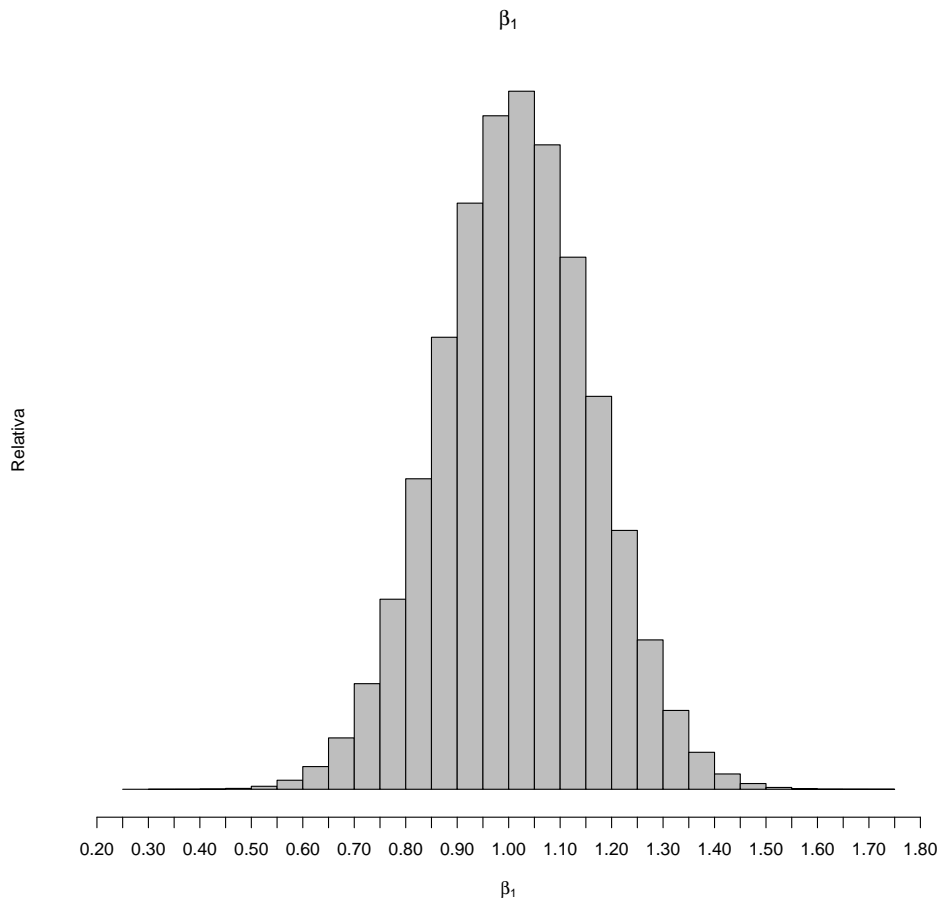


Figura 5.14. *Distribución a priori marginal simulada para β_1*

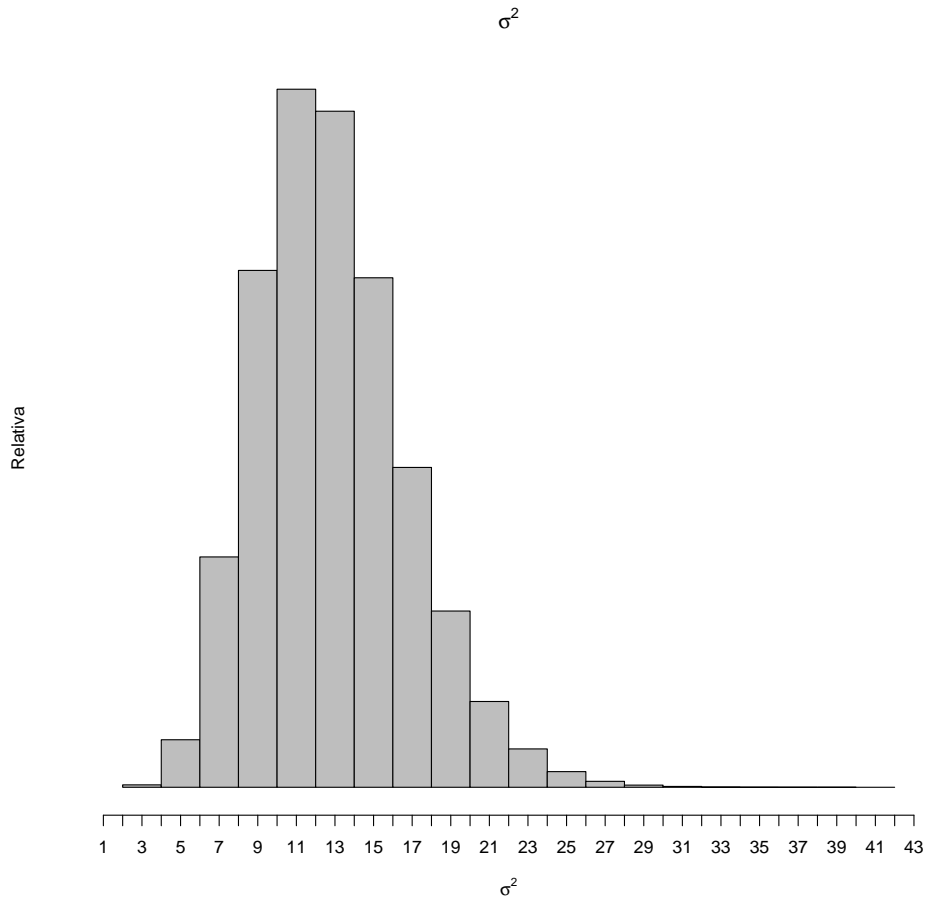


Figura 5.15. *Distribución a priori marginal simulada para σ*

Prueba de normalidad bivariable para (β_0, β_1)

```

library(mvnTest)

HZ.test(resu[1:10000,1:2])
##           Henze-Zirkler test for Multivariate Normality

## data : resu[1:10000, 1:2]

## HZ           : 0.9660402
## p-value      : 0.2840896

## Result  : Data are multivariate normal (sig.level = 0.05)

```

Prueba de Henze-Zirkler para multinormalidad

$$\begin{aligned}
 T &= \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \exp \left[-\frac{\beta^2}{2} (x_i - x_j)' S^{-1} (x_i - x_j) \right] \\
 &= -\frac{2}{(1 + \beta^2)^{k/2}} \sum_{i=1}^N \exp \left[-\frac{\beta^2}{2(1 + \beta^2)} (x_i - \bar{x})' S^{-1} (x_i - \bar{x}) \right] \\
 &= -\frac{N}{(1 + \beta^2)^{k/2}},
 \end{aligned} \tag{5.44}$$

donde:

$$\beta = \frac{1}{\sqrt{2}} \left[\frac{N(2k + 1)}{4} \right]^{1/(k+4)} \tag{5.45}$$

Cuando $N \rightarrow \infty$, los dos primeros momentos de T :

$$E(T) = 1 - \frac{1}{(1 + 2\beta^2)^{k/2}} \left[1 + \frac{k\beta^2}{(1 + 2\beta^2)} + \frac{k(k + 2)\beta^4}{2(1 + 2\beta^2)^2} \right] \tag{5.46}$$

$$\begin{aligned}
 Var(T) &= \frac{2}{(1 + 4\beta^2)^{k/2}} \\
 &+ \frac{2}{(1 + 2\beta^2)^k} \left[1 + \frac{2k\beta^4}{(1 + 2\beta^2)^2} + \frac{3k(k + 2)\beta^8}{4(1 + 2\beta^2)^4} \right] \\
 &- \frac{4}{w^{k/2}} \left[1 + \frac{3k\beta^4}{2w} + \frac{k(k + 2)\beta^8}{2w^2} \right],
 \end{aligned} \tag{5.47}$$

donde:

$$w = (1 + \beta^2) (1 + 3\beta^2) \tag{5.48}$$

- Henze y Zirkler probaron, vía simulación, que T sigue aproximadamente una distribución lognormal.
- Para calcular el *valor - p*, ellos usan una aproximación a la normal:

$$\begin{aligned}
 \kappa &= \log \left(1 + \frac{Var(T)}{E(T)^2} \right) \\
 \nu &= \log(E(T)) - \kappa/2 \\
 Z &= \frac{\log(T) - \nu}{\sqrt{\kappa}} \\
 Valor - p &= 2\Phi(-|Z|)
 \end{aligned} \tag{5.49}$$

Verificamos si la distribución marginal a priori simulada de σ^2 se acomoda a una distribución gamma

```
library(EnvStats)

gofTest(resu[1:5000,3],distribution='gamma')

## Results of Goodness-of-Fit Test
## -----
## Test Method:                Shapiro-Wilk GOF Based on
##                               Chen & Balakrisnan (1995)

## Hypothesized Distribution:    Gamma
## Estimated Parameter(s):      shape = 10.997980
##                               scale = 1.162984

## Estimation Method:          MLE
## Data:                        resu[1:5000, 3]
## Sample Size:                 5000
## Test Statistic:              W = 0.9996789
## Test Statistic Parameter:    n = 5000
## P-value:                     0.6425488
```

Verificamos si la distribución marginal a priori simulada de τ se acomoda a una distribución gamma

```
library(EnvStats)

gofTest(1/resu[1:5000,3],distribution='gamma')

## Results of Goodness-of-Fit Test
## -----
## Test Method:                Shapiro-Wilk GOF Based on
##                               Chen & Balakrisnan (1995)

## Hypothesized Distribution:    Gamma
## Estimated Parameter(s):      shape = 10.346388426
##                               scale = 0.008311766

## Estimation Method:          MLE
## Data:                        1/resu[1:5000, 3]
## Sample Size:                 5000
## Test Statistic:              W = 0.9784269
## Test Statistic Parameter:    n = 5000
## P-value:                     0
## Alternative Hypothesis:      True cdf does not equal the
##                               Gamma Distribution.
```

5.6. Elicitación en el modelo lineal generalizado

El modelo lineal clásico cae en una clase mayor de modelos que se conoce como *modelo lineal generalizado*, *MLG*, la cual tiene tres componentes básicos:

1. Un conjunto de variables aleatorias independientes que pertenecen a la familia exponencial.
2. Una matriz de diseño y un vector de parámetros.
3. Una función link (enlace, conexión) que relaciona las medias del modelo lineal.

Dentro de la clase de modelos lineales generalizados tenemos el modelo lineal clásico, el modelo loglineal, la regresión de Poisson, la regresión logística, etc.

5.6.1. Componentes del *MLG*

En esta parte discutimos de una forma más detallada las componentes del modelo lineal generalizado. Necesitamos, además, revisar ciertos elementos básicos de inferencia estadística, en especial la estimación por máxima verosimilitud.

Las tres componentes del *MLG*

Sean y_1, \dots, y_n variables aleatorias con función densidad de probabilidad (FDP) para y_i , $i = 1, \dots, n$, dada por:

$$f(y_i; \theta_i) = \exp(y_i b_i(\theta_i) + c_i(\theta_i) + d_i(y_i)) \quad (5.50)$$

Las distribuciones de los y_i 's son de la misma forma (todas normales o todas binomiales, etc.), así que las distribuciones de b , c y d no se necesitan.

La FDP conjunta de y_1, \dots, y_n es:

$$f(y_1, \dots, y_n; \theta_1, \dots, \theta_n) = \exp\left(\sum_{i=1}^n y_i b(\theta_i) + \sum_{i=1}^n c(\theta_i) + \sum_{i=1}^n d(y_i)\right) \quad (5.51)$$

y, usualmente, los parámetros θ_i no son de interés directo.

Consideremos un conjunto de parámetros más pequeño β_1, \dots, β_p ($p < n$) tal que la combinación lineal de los β_j es igual a una función del valor esperado de y_i , digamos μ_i ,

$$g(\mu_i) = \mathbf{x}_i^T \beta, \quad (5.52)$$

donde:

- g : link, función monótona y diferenciable
- \mathbf{x}_i : vector de $p \times 1$ variables explicativas
- β : vector de $p \times 1$ de parámetros

El modelo lineal generalizado tiene tres componentes:

1. Variables respuesta y_1, \dots, y_n (de la familia exponencial).
2. Un conjunto de parámetros β y variables explicativas.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} \quad (5.53)$$

3. Una función link g monótona tal que:

$$g(\mu_i) = \mathbf{x}_i^T \beta, \quad (5.54)$$

donde:

$$\mu_i = E[y_i] \quad (5.55)$$

Chen et al. (2003) presentan algunas discusiones sobre la elicitación en el modelo lineal generalizado, pero desde el punto de vista aplicado no se presenta ningún procedimiento que sea utilizable.

5.6.2. Elicitación en el modelo logístico

El modelo logístico es uno de los más utilizados en el trabajo aplicado. La respuesta binaria aparece en muchos problemas de interés en diversas áreas que manejan riesgo, por ejemplo, medicina, finanzas e ingeniería. Bedrick et al. (1997) presentan un resumen de diversos procedimientos para la elicitación de la a priori en el caso de la regresión logística. O'Leary et al. (2009b) presentan tres métodos para la elicitación en el caso de la regresión logística. Entre ellas se encuentra la elicitación de la probabilidad de éxito a diferentes niveles de las covariables. Al-Awadhi y Garthwaite (2006) presentan una metodología de elicitación del modelo logístico en el área de ecología.

5.6.3. Uso de muestras hipotéticas

Vamos a ilustrar esta aproximación para el modelo logístico con una sola variable explicativa.

- Determine los puntos de diseño x_1, \dots, x_k tal que cubramos los puntos de inflexión del modelo.
- Para cada punto de diseño, x_j , ajuste, una beta a priori con parámetros α_j y β_j que representen el conocimiento del experto sobre la probabilidad de éxito.
- De cada distribución *beta* (α_i, β_i) genere un π_i , $i = 1, \dots, k$.

- Para cada punto de diseño x_i , genere una muestra aleatoria de tamaño n_i de éxitos y fracasos con probabilidad de éxito π_i obtenida en el punto anterior, $i = 1, \dots, k$, donde $n^* = \sum_{i=1}^k n_i$ representa el nivel de seguridad que tenga el experto en su conocimiento.
- Con la información así obtenida, estime el modelo logístico y conserve los parámetros estimados del modelo.
- Repita los pasos 3-5 N . Con los N pares de valores hallados asociados con los coeficientes del modelo, ajuste una distribución normal bivariable. Esta será la distribución a priori del experto. El ajuste normal se puede validar mediante el de una estimación de densidad bivariable.
- Un problema a determinar es el número de puntos k en los que debemos observar la variable explicativa.
- Obviamente, debemos escoger puntos que permitan realizar una estimación buena. Un número que podemos escoger es 5.
- Si se escoge un k grande, el proceso puede volverse muy desgastante para el experto y terminar dando respuestas en forma automática.

5.6.4. Determinación de la edad de la menarquia

- La edad de la menarquia ha sido estudiada en muchas poblaciones del mundo y el modelo logístico se utiliza comúnmente para esto. Con este ejemplo se ilustra una forma de conducir una entrevista que nos permita determinar la distribución a priori de los parámetros de un modelo logístico.
- Deseamos determinar los valores plausibles del modelo para ajustarlo a la ciudad de Medellín. Para esto contactamos a un experto en el área cuyo nombre nos reservamos.
- El diálogo ha sido editado y solo algunas parte relevantes son mostradas.
- Estas elicitaciones, en general, tardan varios días y pueden parecer muy repetitivas algunas partes del proceso.
 - *Entrevistador*: Buenas tardes. Como usted es una persona reconocida en el área de crecimiento humano y ha accedido a colaborar, le agradecemos de antemano esta ayuda.
 - *Experto*: No hay problema.
 - *Entrevistador*: ¿Ha escuchado acerca de la modelación matemática de la menarquia a partir de la edad?
 - *Experto*: No. Realmente en la literatura existen valores de referencia, pero las decisiones que se toman involucran mucha información, en general.

- *Entrevistador:* ¿Pero podría darme un valor en el cual considerara que si una niña ha tenido, sin información adicional, la menarquia, sería imposible de aceptar?
- *Experto:* Sí, claro. Nueve años sería esa edad.
- *Entrevistador:* ¿Y cuál sería la edad mínima para la cual una niña normal y sin ninguna información adicional que no hubiese tenido la menarquia, su caso ya fuera considerado un problema médico?
- *Experto:* Yo diría que 16 años y medio.
- *Entrevistador:* El proceso que vamos a comenzar está basado en algo conocido como muestras hipotéticas. Las preguntas que le voy a formular solo pretenden recoger sus impresiones y no tienen necesariamente una respuesta única y verdadera. Es más, puede cambiar la respuesta cuantas veces quiera.
- *Experto:* Está bien.
- *Entrevistador:* Si considera solo la población de las niñas que acaban de cumplir 10 años (no las de 10 años y medio ni las que están cercanas a once, etc.) y tuviera una muestra de 1000 niñas de esta población, ¿cuál sería el número máximo de niñas que usted aceptaría y no se sorprendería, que hubiese tenido la menarquia?
- *Experto:* Bueno, déjeme pensar. En 1000 niñas de esta edad cerrada, quizá 3 o menos.
- *Entrevistador:* Si le dicen que 10 niñas, de este grupo, han tenido la menarquia, ¿se sorprendería?
- *Experto:* Sí. A lo más 5 o menos.
- *Entrevistador:* Y si le dicen que ninguna ha tenido la menarquia, ¿lo aceptaría?
- *Experto:* Sí, es ese número que esperaríamos.
- *Entrevistador:* Pensemos ahora en las quinceañeras. De una muestra de 100 niñas que hayan acabado de cumplir sus quince años, ¿cuál sería el número que esperaríamos encontrar que no haya tenido la menarquia?
- *Experto:* Creo que sería alrededor de 25.
- *Entrevistador:* Y ¿cuál sería, en esta muestra de 100, el número máximo de niñas que no hubiesen tenido la menarquia que no le llevara a decir que esos datos fueran inaceptables?
- *Experto:* Creo que si me dicen que de las 100 ninguna ha tenido la menarquia, no lo aceptaría.
- *Entrevistador:* Sí, pero eso es muy extremo, ¿no es así?

- *Experto*: Sí. Bueno, creo que si me dicen que 30 no han tenido la menarquia, lo tendría como límite.
- *Entrevistador*: Y en la muestra de 100 niñas, yéndonos al otro caso, ¿se sorprendería de que todas hubiesen tenido la menarquia?
- *Experto*: Sí, pero como van las cosas.
- *Entrevistador*: Volvamos a la pregunta, ¿se sorprendería de que todas hubiesen tenido la menarquia?
- *Experto*: Sí, esperaría que al menos 10 no hubieran tenido la menarquia aún.

Para cada grupo de edad considerado, ajustamos una distribución a priori; en la mayoría de ellas determinamos betas.

- Hay en este caso particular una situación en la que, a priori, sabemos que es imposible observar ciertos resultados; por ejemplo, para niñas muy jóvenes, digamos menores de 9 años, estamos convencidos de que no observaremos a alguna que haya tenido la menarquia y lo contrario pasaría con mujeres mayores de 17 años. En este caso hay seguridad absoluta y el experto tendría una información equivalente a un tamaño muestral muy grande. Pero en lo que realmente estamos interesados es en la determinación del modelo en el que no hay certidumbre plena. Por lo tanto, en nuestra entrevista debemos buscar unos valores de edades en los cuales el experto reconozca esto. Por ejemplo, edades como 11, 12, 13, 14 y 15 años (edades exactas pueden ser buenos valores).

Elicitación de la distribución de β del modelo logístico

Tabla 5.4. *Valores elicitados por grupos de edad*

Edad	10	12	14	16
Número de niñas	200	100	100	100
Número con menarquia	15	20	60	95
Proporción	15/200=0.075	0.20	0.60	0.95
$n_{equivalente}$	5	10	20	20

```

# genera muestra
muestra<-function(propor,edad,n.eq){
  resu<-NA
  for(ii in 1:length(propor))resu<-rbind(resu,
    cbind(sample(c(1,0),n.eq[ii],replace=T,
      prob=c(propor[ii],1-propor[ii])),edad[ii]))
  return(resu)
} # fin muestra

propi<-c(0.075,0.20,0.6,0.95)
neq<-c(5,10,20,20)
edad<-c(10,12,14,16)

betas<-NULL
for(i in 1:1000){
  result<- muestra(propri,edad,neq)
  yy<-result[,1]
  xx<-result[,2]
  tempo<-glm(yy~xx,family=binomial)$coefficients
  betas<-rbind(betas,tempo)
} #fin for

library(robust)
## Loading required package: fit.models
covRob(betas)
## Call:
covRob(data = betas)

## Robust Estimate of Covariance:
##           (Intercept)          xx
## (Intercept)   16.470  -1.15549
## xx             -1.155   0.08185

## Robust Estimate of Location:
## (Intercept)          xx
##      -13.595          1.009

solve(matrix(c(16.470, -1.15549, -1.15549,0.08185),ncol=2))
##           [,1]      [,2]
## [1,]  6.338888  89.48713
## [2,] 89.487128 1275.52207

```

Un problema

No siempre el experto está absolutamente seguro del número de niñas que han tenido la menarquia en un grupo de n para una determinada edad. Es más fácil que dé un rango de valores con alta seguridad y en ese intervalo el experto pueda escoger el valor más creíble.

Tabla 5.5. *Estimados vía elicitación por grupos de edad*

Edad	10	12	14	16
Número de niñas	200	100	100	100
Mínimo número aceptable	5	14	50	93
Número más creíble con menarquia	15	20	60	95
Máximo número aceptable	20	22	66	98
Proporción	15/200=0.075	0.20	0.60	0.95
Proporción mínima	5/200=0.025	0.14	0.50	0.93
Proporción máxima	20/200=0.10	0.22	0.66	0.98
$n_{equivalente}$	5	10	20	20

5.7. Elicitación de la distribución a priori informativa del coeficiente de correlación

El coeficiente de correlación es una de las medidas estadísticas de más uso en el trabajo aplicado. Algunas de sus propiedades fueron estudiadas por Zheng y Matis (1993). Una discusión sobre sus interpretaciones puede hallarse en Falk y Well (1997). La estimación del coeficiente de correlación por medio de intervalos es importante y para esto se dispone de diversos métodos. El problema para el analista es la carencia de reglas sobre cuál fórmula es preferible. Para esto hemos realizado un estudio de simulación que nos permite analizar el comportamiento de los niveles de confianza reales y compararlo con los teóricos de los diversos intervalos disponibles.

Asumamos que $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ es una muestra aleatoria de una normal bivariable con vector de medias $\boldsymbol{\mu}$ y matriz de varianzas y covarianzas $\boldsymbol{\Sigma}$. El estimador máximo verosímil de ρ es (Graybill, 1976):

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2\right)^{1/2}} \quad (5.56)$$

y el estimador UMVUE (estimador insesgado de varianza uniformemente mínima) de ρ es:

$$\hat{\rho} = R \left(\frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n-3}{2}\right)} \right) \int_0^1 \frac{t^{-1/2} (1-t)^{(n-5)/2}}{\sqrt{1-t(1-R^2)}} dt \quad (5.57)$$

La FDP de R es:

$$f_R(r) = \frac{(n-2)(1-\rho^2)^{(n-1)/2}}{\pi} (1-r^2)^{(n-4)/2} \int_0^\infty (\cosh w - \rho r)^{-(n-1)} dw, \quad (5.58)$$

donde $-1 < r < 1$ y $-1 < \rho < 1$. El único parámetro de la distribución es ρ .

Si $\pi(\rho)$ denota la distribución a priori de ρ , entonces, la distribución a posteriori será:

$$\pi(\rho; \text{Datos}) \propto L(\rho; \text{Datos})\pi(\rho) \quad (5.59)$$

$$\pi(\rho; \text{Datos}) \propto \pi(\rho) (1 - \rho^2)^{(n-1)/2} \int_0^\infty (\cosh w - \rho r)^{-(n-1)} dw \quad (5.60)$$

Si escogemos una distribución a priori no informativa de Laplace, $\pi(\rho) \propto 1$, entonces, los resultados de la distribución a posteriori, serán equivalentes a los clásicos, ya que esta última será proporcional a la función de verosimilitud.

$$\pi(\rho; \text{Datos}) \propto L(\rho; \text{Datos}) \quad (5.61)$$

Dada la complejidad de la función conjunta, una aproximación que se ha presentado es considerar la distribución conjunta de los datos normales bivariabes $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ con parámetros $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$. La distribución posterior será:

$$\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho | \text{Datos}) \propto L(\text{Datos} | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \quad (5.62)$$

Una ventaja de esta aproximación es que existen resultados sobre distribuciones conjugadas para el caso normal, pero el problema es que exige la elicitación de cuatro parámetros adicionales, los cuales no son de interés del investigador.

En este caso, los parámetros $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ son considerados como parámetros de perturbación (nuisance) y hay que eliminarlos vía marginalización.

$$\pi(\rho | \text{Datos}) = \int \int \int \int \pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho | \text{Datos}) d\mu_1 d\mu_2 d\sigma_1^2 d\sigma_2^2. \quad (5.63)$$

5.7.1. Conclusiones y recomendaciones

Al ser la elicitación un proceso extremadamente complejo e infravalorado, se necesita investigación para comparar científicamente los diferentes métodos que se han propuesto. Se requiere que los psicólogos y los científicos cognitivos comprendan mejor el proceso profundo que tiene lugar cuando un experto se enfrenta a una tarea de obtención particular. Se ha propuesto un algoritmo basado en una muestra mental hipotética para determinar los parámetros de la distribución conjugada de una población normal con parámetros desconocidos. Según nuestra experiencia, los expertos se sienten cómodos cuando se les hacen preguntas sobre los cuantiles de este tipo de muestra hipotética. Es fácil generalizar el algoritmo a otras situaciones, incluidos los modelos más complejos, como los de regresión.

Capítulo 6

Elicitación a partir de información disponible de un solo estudio

Muchas veces la información disponible sobre los parámetros de interés puede ser utilizada para la construcción de distribuciones a priori. Una idea que es general es que este proceso es directo y fácil. Esto es equivocado y los elicidores son conscientes de que el conocimiento de la información disponible, sus antecedentes y su calidad son fundamentales. Se puede disponer de información directa sobre parámetros que es posible extender a nuestro caso sin mayor problema. En otras ocasiones, cuando los estudios hacen referencia al mismo parámetro, pero provienen de épocas o de lugares diferentes, es necesario ajustar las distribuciones a priori de tal forma que se acomoden a nuestra situación particular. Por ejemplo, si elicítamos la estatura promedio de hombres adultos jóvenes en Medellín y tenemos información sobre esta variable realizada en Bogotá hace muy poco, se podría construir la a priori casi directamente, mientras que si el estudio se realizó en Perú, debemos efectuar ajustes que permitan considerar las posibles diferencias en las poblaciones o nuestra desconfianza en esta a priori (usualmente se modifican las varianzas).

6.1. Uso de estudios previos para la determinación de la distribución a priori

La determinación de la distribución a priori utilizando información previa puede facilitar el proceso de elicitación. Si el experto considera que esta información proviene de situaciones similares a la que se tiene, es posible construir las a priori casi en forma directa. Es más, el conocimiento a priori del investigador está altamente influenciado por esto, (Kennedy, 1986).

Si se dispone de varios estudios, es posible realizar un metaanálisis para obtener una a priori. Un problema ocurre cuando la información de otros estudios apenas es comparable o no hay una justificación clara para asumir que lo es. En este caso, el investigador debe usar su conocimiento personal para ajustar los valores hallados así.

Papathomas y Hocking (2003) discuten el proceso de elicitación del riesgo de reparación o remplazo de partes en la industria del agua en el Reino Unido. Ellos presentan el proceso de la determinación de la distribución a priori conjunta a partir de distribuciones a priori marginales mediante el uso de cópulas.

6.2. Elicitación de los parámetros de una a priori cuasiconjugada para una población

Si se dispone de información publicada sobre los parámetros de una población, usualmente se presenta un intervalo de confianza junto con el nivel de confianza y el tamaño muestral; podemos usar un programa como el siguiente para determinar la a priori conjunta de la media y la precisión.

```

Construye.apriori.Normal.IC<-function(LI,LS,n,nivel){
  media<-(LI+LS)/2
  alfa<-1-nivel
  alfa.2<-alfa/2
  gl<-n-1
  t.teo<-qt(1-alfa.2,gl)
  S2<-(LS-media)^2*n/t.teo^2
  precision<-n/S2
  LI.var<-gl*S2/qchisq(1-alfa.2,gl)
  LS.var<-gl*S2/qchisq(alfa.2,gl)
  LI.preci<-1/LS.var
  LS.preci<-1/LI.var

  a.mini<-function(parame,LI=LI,LS=LS,nivel=nivel){
    a0<-parame[1]
    b0<-parame[2]
    alfa<-1-nivel
    alfa.2<-alfa/2
    res<-(LI-qgamma(alfa.2,a0,rate=b0))^2+(LS-qgamma(1-alfa.2,a0,rate=
      b0))^2
    return(res)
  }
  res<-optim(c(1,1),a.mini,method="L-BFGS-B",lower=c(0.0001,0.0001),
    upper=c(Inf,Inf),LI=LI,LS=LS,nivel=nivel)
  alfa0<-res$par[1]
  beta0<-res$par[2]
  list(media=media,precision=precision,alfa0=alfa0,beta0=beta0)
}

Construye.apriori.Normal.IC(10,20,20,0.95)

```

6.3. Elicitación de una proporción

Si deseamos estudiar la proporción de niñas que a los doce años han tenido la menarquia en Medellín, y el año inmediatamente anterior se realizó un estudio similar (en la misma ciudad y con la misma población objetivo), podríamos utilizar las distribuciones muestrales casi en forma directa. Sin embargo, si el estudio se adelantó en una ciudad norteamericana, es posible usar la información, pero es preferible incrementar la varianza de las distribuciones muestrales, de tal forma que le restemos confianza a esta información.

Por ejemplo, si en el estudio previo obtuvimos un intervalo de confianza del 95 % para esta probabilidad 0.08 y 0.14 obtenemos $\hat{\pi} = (0.08 + 0.14)/2 = 0.11$ y $n = 418$. La varianza del estimador será 0.0002342105. Si ajustamos esta información a una a priori beta, obtenemos los parámetros a priori de la siguiente forma:

$$Media = 0.11 = \frac{\alpha}{\alpha + \beta} \quad (6.1)$$

$$Varianza = \frac{0.11(1 - 0.11)}{418} = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}, \quad (6.2)$$

de donde obtenemos:

$$\begin{aligned} \alpha_0 &= 45.87 \\ \beta_0 &= 371.13 \end{aligned}$$

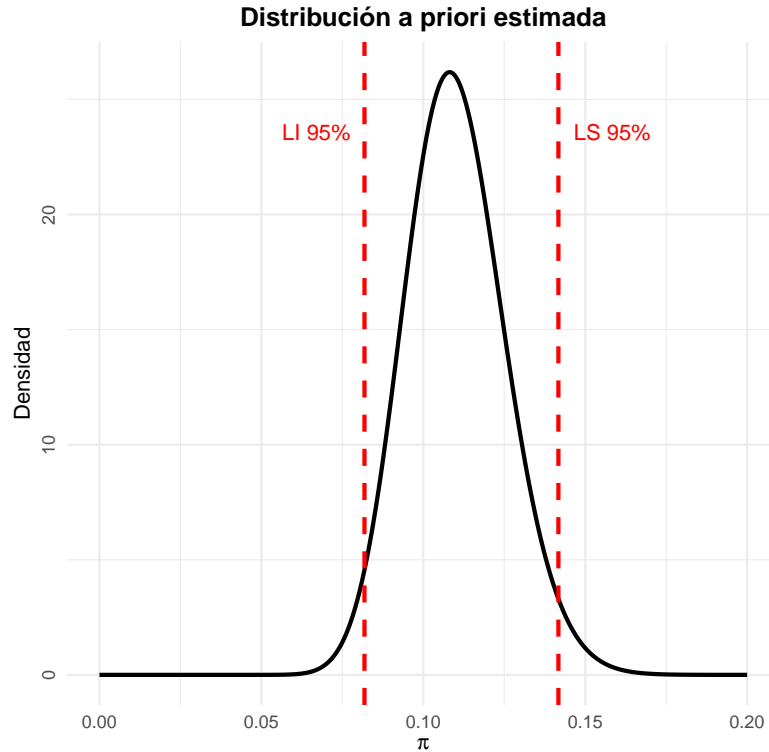


Figura 6.1. Intervalo de credibilidad y distribución a priori elicitada para la proporción

6.4. Elicitación de los parámetros de una distribución multinomial

La distribución multinomial juega un papel fundamental en el trabajo aplicado, siendo la generalización multivariable de la distribución binomial.

Definición 1 (distribución Dirichlet). *El vector aleatorio $\mathbf{X} = (X_1, \dots, X_k)'$ se distribuye como una Dirichlet con vector de parámetros $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)'$ con $\alpha_i > 0$; $i = 1, \dots, k$, si la p.d.f. $f(\mathbf{x}|\boldsymbol{\alpha})$ para $\mathbf{x} = (x_1, \dots, x_k)$ y $\sum_{i=1}^k x_i = 1$ está dada por:*

$$f(\mathbf{x}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} x_1^{\alpha_1-1} \dots x_k^{\alpha_k-1} \quad (6.3)$$

La media de X_i es:

$$E(X_i) = \frac{\alpha_i}{\alpha_0}, \quad (6.4)$$

donde $\alpha_0 = \sum_{i=1}^k \alpha_i$.

La varianza de X_i es:

$$\text{var}(X_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}, \quad (6.5)$$

y la covarianza entre X_i y X_j es, ($i \neq j$),

$$\text{Cov}(X_i, X_j) = -\frac{\alpha_i\alpha_j}{\alpha_0^2(\alpha_0 + 1)} \quad (6.6)$$

Teorema 6.1. *Suponga que $\mathbf{Y} = (Y_1, \dots, Y_k)'$ tiene una distribución multinomial con parámetros n (fijo) y $\mathbf{W} = (W_1, \dots, W_k)'$, desconocidos. Suponga también que la distribución a priori de \mathbf{W} es una Dirichlet con vector de parámetros $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)'$ con $\alpha_i > 0$; $i = 1, \dots, k$. Entonces, la distribución posterior de \mathbf{W} cuando $Y_i = y_i$, $i = 1, \dots, k$, es una distribución Dirichlet con vector de parámetros $\boldsymbol{\alpha}^* = (\alpha_1 + y_1, \dots, \alpha_k + y_k)'$.*

El parámetro α_k puede ser interpretado como el conteo a priori, antes de ver los datos, que esperaríamos ver en la celda k . Un valor grande para este parámetro muestra un gran conocimiento previo acerca de la distribución, mientras que valores pequeños corresponden a poco conocimiento.

Flórez-Rivera y Correa-Morales (2015) presentan una propuesta de elicitación de los parámetros. Gupta y Upadyay (2019) presentan una propuesta de elicitación de los parámetros de una Dirichlet usando una muestra disponible. Flórez-Rivera et al. (2015) presentan un programa en R para su elicitación. Montoya y Correa (2017) presentan un procedimiento de elicitación de los parámetros de la distribución multinomial cuando se utilizan varios expertos.

Tipos de sangre. La siguiente tabla presenta los datos sobre el tipo de sangre en una muestra de personas de las regiones central y oriental de Antioquia.

Tabla 6.1. *Distribución de frecuencias de los tipos de sangre*

	Tipo de sangre			
	O	A	AB	B
Frecuencia	474	246	11	59

Si no tenemos un conocimiento a priori sobre las diversas proporciones, digamos π_O, π_A, π_{AB} y π_{AB} , entonces podemos escoger como a priori una $Dirichlet(1, 1, 1, 1)$. Entonces la a posteriori será $Dirichlet(474 + 1, 246 + 1, 11 + 1, 59 + 1)$.

El proceso de elicitación de los parámetros de una multinomial usando datos pasados puede hacerse de dos formas:

1. Elicitar parámetro por parámetro, similar a lo que se hace con el caso de una sola proporción, y en este caso obtenemos una a priori que puede aproximarse como el producto de betas.
2. Puede elicitar mediante modelos jerárquicos, en los cuales se consideran unas hiperdistribuciones que generan los parámetros de una Dirichlet, que es la que finalmente controla el proceso estocástico (Seo y Kim, 2019). Ellos lo que exponen es de carácter general utilizando MCMC, pero no lo hacen con la forma de elicitar hiperparámetros. Vamos usando los datos que se tienen sobre los tipos de sangre para ilustrar este proceso. El algoritmo sería:
 - a) Con la muestra original calcule el vector estimado de proporciones, digamos $\hat{\boldsymbol{\pi}}$.
 - b) Genere una muestra de tamaño n^* de esta población y estime $\hat{\boldsymbol{\pi}}^{(j)}$, $j = 1, \dots, N$. El n^* refleja el nivel de confianza en los datos a la mano y debe considerar la pertinencia para el problema concreto que tenemos.
 - c) Para cada muestra así calculada aplique el teorema anterior y obtenga $\boldsymbol{\alpha}$, el vector de parámetros de una Dirichlet. Así obtenemos $\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(N)}$.
 - d) Para cada componente del vector de parámetros de la Dirichlet ajuste una hiperdistribución *gamma* (a_i, b_i), $i = 1, \dots, k$, donde (a_i, b_i) es el estimador obtenido usando $\alpha_1^{(i)}, \dots, \alpha_N^{(i)}$.

6.4.1. Problema de un torneo de tenis

Los torneos de tenis tienen características que pueden ayudar a construir la distribución a priori, vía simulación, de la probabilidad de que cada uno de los participantes gane el torneo, con cierta facilidad (esto aplica a torneos como el mundial de fútbol):

- La ATP tiene una clasificación mediante puntos de los tenistas profesionales (usualmente los jugadores reciben puntos que varían de acuerdo con la importancia de los torneos en los que participan, de la calidad de los rivales enfrentados y si ganan o no ante estos rivales).
- En un torneo típico, a los jugadores se les crea un camino de tal forma que el mejor ranqueado juega contra el peor, el segundo mejor contra el segundo peor, etc.
- La probabilidad de que el jugador A le gane al jugador B (ambos con puntaje de la ATP, p_A y p_B , respectivamente), será aproximada mediante:

$$P(\text{gane A}) = \frac{p_A}{p_A + p_B} \quad (6.7)$$

- Jugador que pierde, jugador que se elimina del torneo.

```

puntos<-c(800,780,600,500,490,200,150,100)
puestos<-order(puntos,decreasing=T)

# Argumento x: vector con 4 componentes
# x[1]: rango del mejor de los jugadores
# x[2]: puntaje del mejor jugador
# x[3]: rango del otro jugador,
# x[4]: puntaje del otro jugador

un.partido<-function(x)sample(c(x[1],x[3]),1,prob=c(x[2],x[4]))

# Arg. aaaa: argumento dummy
torneo<-function(aaaa,puntos,puestos){
  n<-length(puntos)
  ciclos<-log2(n)
  # En cada etapa se eliminan la mitad de los jugadores
  datos<-cbind(puestos,puntos)
  for(i in ciclos:1){
    n<-2^i
    n.m<-n/2
    partidos<-matrix(cbind(datos[1:n.m,],datos[n:(n.m+1),]),nrow=n.m)
    res<-apply(partidos,1,un.partido)
    datos<-datos[datos[,1] %in% res,]
  } #Fin del for
  return(res)
} # Fin función torneo

# Torneos ganados por cada jugador en 1000 torneos
(res<-table(apply(matrix(1,nrow=1000,ncol=1),1,torneo,
  puntos,puestos)))

##   1   2   3   4   5   6   7   8
## 326 300 183  96  81  7  6  1

# Probabilidad simulada para cada jugador
(proba.est<-res/sum(res))

##   1   2   3   4   5   6   7   8
## 0.326 0.300 0.183 0.096 0.081 0.007 0.006 0.001

```

Dado que estos resultados son estimadores puntuales, podemos considerar la aproximación a la media y a la varianza de cada una de las 8 probabilidades como:

$$E(\pi_i) \approx \pi_{i,Simulado} \quad (6.8)$$

$$Var(\pi_i) \approx \frac{\pi_{i,Simulado}(1 - \pi_{i,Simulado})}{n_{equivalente}}, \quad (6.9)$$

donde el $n_{equivalente}$ nos mide el nivel de confianza en estos resultados.

Si asumimos como una a priori una distribución Dirichlet con parámetro α , entonces podemos proceder así:

1. Con las probabilidades estimadas del proceso de simulación, digamos $\pi_{Simul} = (0.326, 0.300, 0.183, 0.096, 0.081, 0.007, 0.006, 0.001)$, generamos N muestras de esta multinomial con tamaño igual a nuestro nivel de seguridad, digamos n_{equ} .
2. A partir de los N vectores del paso anterior se estiman los parámetros de la distribución Dirichlet así:

$$\alpha_j = \sum_i \frac{n_j^{(i)}}{N} + 1 \quad (6.10)$$

```
proba<-scan()
0.326 0.300 0.183 0.096 0.081 0.007 0.006 0.001

replicas<-rmultinom(1000,20,proba)
alfas<-apply(replicas,1,mean)+1
alfas
## [1] 7.660 6.953 4.630 2.880 2.601 1.133 1.116 1.027
alfas/sum(alfas)
## [1] 0.27357 0.24832 0.16535 0.10285 0.09289 0.04046 0.03985
## [8] 0.03667
```

6.5. Elicitación del parámetro de un modelo de Poisson

La siguiente tabla presenta la distribución del número de goles marcados por el equipo local en el torneo de fútbol colombiano.

Tabla 6.2. *Distribución de frecuencias de los tipos de sangre*

Goles	0	1	2	3	4	5	6 o más
Número de juegos							

Si deseamos construir la a priori conjugada para el parámetro, la distribución $gamma(\alpha, \beta)$, podemos proceder así:

1. Con la información anterior calculamos una aproximación a la distribución muestral del estimador clásico de λ que es la media muestral, digamos $\hat{\lambda} \sim N(\lambda; \lambda/n)$. Usamos a $\hat{\lambda}$ como estimador de λ , o se, la distribución a considerar $N(\hat{\lambda}; \hat{\lambda}/n)$.
2. Generamos al azar N (grande) valores de la distribución $N(\hat{\lambda}; \hat{\lambda}/n)$, digamos $\lambda_1, \dots, \lambda_N$.

3. Para cada λ_i , $i = 1, \dots, N$, generamos una muestra simulada de la distribución de $Poisson(\lambda_i)$ de tamaño n^* , donde este tamaño depende del nivel de confianza en la calidad y pertinencia de la información muestral.
4. Para cada muestra estimamos λ_i , para $i = 1, \dots, N$. A estos valores le ajustamos la distribución $gamma$. Los valores de α y β serán los parámetros de la a priori deseados.

Observe que hay un proceso en el que el experto debe decidir cuál es su n^* equivalente a su nivel de conocimiento.

6.6. Elicitación de los parámetros del modelo lineal

Cuando se dispone de información previa sobre los parámetros de un modelo lineal, podemos tenerla en forma de un modelo estimado junto con sus errores estimados o de los datos originales. Debemos tener siempre en mente que esta información no corresponde verdaderamente a la población a la que nos enfrentamos, sino a otra que puede ser más o menos parecida. Dependiendo de esto, necesitamos realizar ajustes que incrementen el nivel de entropía en las distribuciones a priori elicítadas de esta información.

Para ilustrar el caso del modelo lineal supongamos que deseamos modelar el precio de oferta de un automóvil Twingo hoy y disponemos del modelo calculado a partir de la información aparecida en el periódico El Colombiano el 4 de marzo de 2012 (esta forma de presentar resúmenes de modelos lineales es bastante común en revistas especializadas y, en general, no se entrega información de la matriz de covarianzas estimada entre los estimadores de los coeficientes, lo cual es una lástima).

Tabla 6.3. *Parámetros estimados del modelo*

	Estimación	Error típico	valor t	$\Pr(> t)$
Intercepto	2.0982	0.9938	2.11	0.0463
año	0.7295	0.0676	10.80	0.0000
$\hat{\sigma}_e = 1.309$		g.l.=22	$R^2 = 0.8412$	

Una forma de generar una distribución a priori conjunta del vector de parámetros del modelo (β_0, β_1, τ) , donde $\tau = 1/\sigma_e^2$, la precisión es:

1. $\tau \sim gamma(\alpha_0, \beta_0)$ donde usamos como $E(\tau) = 1/(1.309)^2 = \alpha_0/\beta_0$. Recordando que $Var(\tau) = \alpha_0/\beta_0^2$, podemos usar el intervalo de confianza que se construya para σ_e^2 ; el del 95% es (1.024901; 3.432479). Podemos, por lo tanto, considerar que $(1/3.432479; 1/1.024901) = (0.2913346; 0.975704)$ es un intervalo aproximado para τ . El $\alpha_0 = 0.7963813$ es hallado vía numérica usando:

```
f.a.optim<-function(a0)(pgamma( 0.2913346 , a0 , a0/1.309) -0.025) ^2
                        +(pgamma( 0.975704 , a0 , a0/1.309) -0.975) ^2
optim(1 , f.a.optim , lower=0.000001 , method='L-BFGS-B')
```

teniendo en cuenta que $\beta_0 = \alpha_0/1.309$. Por lo tanto, $\beta_0 = 0.6083891$.

2. $\beta_0 \sim N(2.0982, k_1 0.9938^2)$
3. $\beta_1 \sim N(0.7295, k_2 0.0676^2)$,

donde k_1 y k_2 son factores que se escogen de tal forma que nos inflen estas varianzas de modo que involucren consideraciones como la pertinencia del modelo para nuestro caso.

La a priori inicial será

$$\xi(\beta_0, \beta_1, \tau) = \xi(\beta_0) \xi(\beta_1) \xi(\tau) \quad (6.11)$$

6.7. Elicitación de los parámetros de una normal multivariable

El análisis multivariable está conformado por un conjunto de técnicas para analizar simultáneamente más de una variable. Usualmente, se limita al análisis de variables continuas con el supuesto de que provienen de una distribución normal multivariable. Al-Awadhi y Garthwaite (2001) revisan y comparan métodos para elicitar los parámetros de esta distribución. Al-Alwadhi y Garthwaite (1998) presentan un procedimiento para elicitar los parámetros de la normal multivariable. Al-Alwadhi y Garthwaite (2006) presentan una aplicación de estos procedimientos. Elicitación de distribuciones multivariables en forma no paramétrica ha sido propuesta por Moala y O'Hagan (2010).

6.7.1. Distribución normal multivariable

Un modelo matemático para representar la relación entre un conjunto de k variables aleatorias es la normal multivariable. Si \mathbf{x} es un vector aleatorio con función de densidad de probabilidad:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad (6.12)$$

decimos que $\mathbf{x} \sim N(\boldsymbol{\mu}, \Sigma)$, donde $\boldsymbol{\mu}$ es el vector de medias de la distribución y Σ es la matriz de varianzas y covarianzas de la población.

6.7.2. Distribuciones condicionales

Si $\mathbf{y} = (\mathbf{y}'_1, \mathbf{y}'_2) \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces \mathbf{y}_1 y $\mathbf{y}_{2.1} = \mathbf{y}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{y}_1$ tienen las siguientes distribuciones y, además, son estadísticamente independientes:

$$\mathbf{y}_1 \sim N_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}) \quad (6.13)$$

$$\mathbf{y}_{2.1} \sim N_s(\boldsymbol{\mu}_{2.1}, \boldsymbol{\Sigma}_{22.1}), \quad (6.14)$$

donde:

$$\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\mu}_1 \quad (6.15)$$

$$\boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \quad (6.16)$$

Usando los supuestos del resultado anterior, la distribución condicional de \mathbf{y}_2 dado \mathbf{y}_1 es:

$$\mathbf{y}_2 | \mathbf{y}_1 \sim N_s(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22.1}) \quad (6.17)$$

6.7.3. Distribución de Wishart

Definición: decimos que la matriz aleatoria simétrica y definida positiva $\mathbf{C}_{p \times p}$ se distribuye Wishart, si su FDP está definida como:

$$f(\mathbf{C} | \nu, \boldsymbol{\Sigma}) = c |\boldsymbol{\Sigma}|^{-\nu/2} |\mathbf{C}|^{(\nu-p-1)/2} \exp \left\{ -\frac{1}{2} \text{traza}(\boldsymbol{\Sigma}^{-1}\mathbf{C}) \right\}, \quad (6.18)$$

donde:

$$c = \left[2^{\nu p/2} \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma \left(\frac{\nu + 1 - j}{2} \right) \right]^{-1} \quad (6.19)$$

6.7.4. Distribución t multivariable

Definición: decimos que el vector aleatorio \mathbf{x} se distribuye t multivariable, si su FDP está definida como:

$$f(\mathbf{x} | n, \boldsymbol{\mu}, \mathbf{T}) = c \left[1 + \frac{1}{n} (\mathbf{x} - \boldsymbol{\mu})' \mathbf{T} (\mathbf{x} - \boldsymbol{\mu}) \right]^{-(n+k)/2}, \quad (6.20)$$

donde:

$$c = \frac{\Gamma((n+k)/2) |\mathbf{T}|^{1/2}}{\Gamma(n/2) (n\pi)^{k/2}} \quad (6.21)$$

- $E(\mathbf{X}) = \boldsymbol{\mu}$
- $var - cov(\mathbf{X}) = \frac{n}{n-2} \mathbf{T}^{-1}$

Supongamos que $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ es una muestra de una distribución multivariable k -dimensional, $N_k(\mathbf{0}, \boldsymbol{\Sigma})$. Definamos:

$$V = \sum_{i=1}^n \mathbf{Y}_i \mathbf{Y}_i' \sim Wishart(n, \boldsymbol{\Sigma}) \quad (6.22)$$

Resultado:

- Suponga que $\mathbf{X}_1, \dots, \mathbf{X}_n$ es una muestra aleatoria (m.a.) de una $N(\mathbf{m}, \mathbf{R})$, donde \mathbf{m} es desconocida y la precisión \mathbf{R} es conocida.
- Suponga que la a priori de \mathbf{m} es una $N(\boldsymbol{\mu}_0, \boldsymbol{\tau}_0)$. Asumimos que $\boldsymbol{\tau}_0 > 0$.
- La distribución posterior de \mathbf{m} , dados los datos, es $N(\boldsymbol{\mu}_1, \boldsymbol{\tau}_0 + n\mathbf{R})$, donde:

$$\boldsymbol{\mu}_1 = (\boldsymbol{\tau}_0 + n\mathbf{R})^{-1} (\boldsymbol{\tau}_0 \boldsymbol{\mu}_0 + n\mathbf{R} \bar{\mathbf{X}}) \quad (6.23)$$

Resultado:

- Suponga que $\mathbf{X}_1, \dots, \mathbf{X}_n$ es una m.a. de una $N(\mathbf{m}, \mathbf{R})$, donde \mathbf{m} es conocida y la precisión \mathbf{R} es desconocida.
- Suponga que la a priori de \mathbf{R} es una $Wishart(\alpha, \boldsymbol{\tau}_0)$. Asumimos que $\alpha > k - 1$ y $\boldsymbol{\tau}_0 > 0$.
- La distribución posterior de \mathbf{R} , dados los datos, es $Wishart(\alpha + n, \boldsymbol{\tau}_1)$, donde:

$$\boldsymbol{\tau}_1 = \boldsymbol{\tau}_0 + \sum_{i=1}^n (\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})' \quad (6.24)$$

6.7.5. Familia conjugada de los parámetros de la población normal multivariable $N(\mathbf{m}, \mathbf{R})$

Resultado:

- Suponga que $\mathbf{X}_1, \dots, \mathbf{X}_n$ es una m.a. de una $N(\mathbf{m}, \mathbf{R})$, donde \mathbf{m} es desconocida y la precisión \mathbf{R} es desconocida.
- Suponga que la a priori conjunta de \mathbf{m} y \mathbf{R} es como sigue:
 - $\mathbf{m} | \mathbf{R} \sim N(\boldsymbol{\mu}_0, \nu \mathbf{R})$
 - $\mathbf{R} \sim Wishart(\alpha, \boldsymbol{\tau})$. Asumimos que $\alpha > k - 1$ y $\boldsymbol{\tau}_0 > 0$.

- La distribución posterior conjunta de \mathbf{m} y \mathbf{R} , dados los datos, es como sigue:

- $\mathbf{m} | \mathbf{R} \sim N(\boldsymbol{\mu}_1, (\nu + n)\mathbf{R})$, donde:

$$\boldsymbol{\mu}_1 = \frac{\nu\boldsymbol{\mu}_0 + n\bar{\mathbf{x}}}{\nu + n} \quad (6.25)$$

- $\mathbf{R} \sim Wishart(\alpha + n, \boldsymbol{\tau}_1)$, donde:

$$\boldsymbol{\tau}_1 = \boldsymbol{\tau} + \mathbf{S} + \frac{\nu n}{\nu + n} (\boldsymbol{\mu}_0 - \bar{\mathbf{x}})(\boldsymbol{\mu}_0 - \bar{\mathbf{x}})' \quad (6.26)$$

Resultado: distribución marginal de \mathbf{m}

- Suponga que $\mathbf{X}_1, \dots, \mathbf{X}_n$ es una m.a. de una $N(\mathbf{m}, \mathbf{R})$, donde \mathbf{m} es desconocida y la precisión \mathbf{R} es desconocida.
- Suponga que el resultado anterior cumple.
- La distribución posterior marginal de \mathbf{m} , dados los datos, es $t_{multi}(\alpha + n - k + 1, \boldsymbol{\mu}_1, (\nu + n)(\alpha + n - k + 1)\boldsymbol{\tau}_1)$.

```

pies.dat<-scan()
24.2  9.4  5.5  3.0  3.2  1
21.7  8.5  6.1  3.2  2.6  0
25.4  9.6  5.5  4.0  3.1  1
25.0 10.1  5.3  3.5  2.7  1
22.0  8.5  5.7  3.1  2.7  0
25.9  9.3  6.1  4.3  3.3  1
22.2  8.6  5.2  3.9  2.9  0
21.7  8.4  5.0  3.2  2.3  0
25.5  9.2  6.1  3.3  3.2  1
24.4  9.4  4.7  3.6  2.8  1

pies.dat<-matrix(pies.dat,ncol=6,byrow=T)[,-6]
n<-nrow(pies.dat)
S<-cov(pies.dat)
x.b<-colMeans(pies.dat)
# Media a priori
mu.0<-c(1,1,1,1,1)
tau.0<-diag(rep(1,5))
nu.0<-1
alfa.0<-1

# Para a posteriori
mu.1<-(nu.0*mu.0+n*x.b)/(nu.0+n)
tau.1<-tau.0+S+nu.0*n/(nu.0+n)*(mu.0-x.b)%*%t(((mu.0-x.b)))

library(MCMCpack)
library(MASS)
library(hdrcde)
result<-NULL
varianzas<-array(NA,dim=c(nrow(tau.1),ncol(tau.1),1000))

```

```

for(i in 1:1000){
  # Genera R de la Wishart
  R<-rwish(alfa.0+n,tau.1)
  Prec.1<-(nu.0+n)*R
  res<-mvrnorm(1,mu.1,temp<-solve(Prec.1))
  varianzas[,i]<-temp
  result<-rbind(result,res)
}

hdr.den(result[,1],main='Distribución de la media de la longitud',
  ylab='Densidad')
## $hdr
##          [,1]      [,2]
## 99% 21.58753 21.86368
## 95% 21.62359 21.82387
## 50% 21.69686 21.75925

## $mode
## [1] 21.72843

## $falpha
##          1%          5%          50%
## 0.2926026 1.1361226 6.4336162

hdr.den(result[,2],main='Distribución de la media de la anchura Má
  xima',ylab='Densidad')
## $hdr
##          [,1]      [,2]
## 99% 8.022526 8.721809
## 95% 8.139914 8.598297
## 50% 8.291337 8.439612

## $mode
## [1] 8.364561

## $falpha
##          1%          5%          50%
## 0.1234953 0.6606465 2.5470789

```

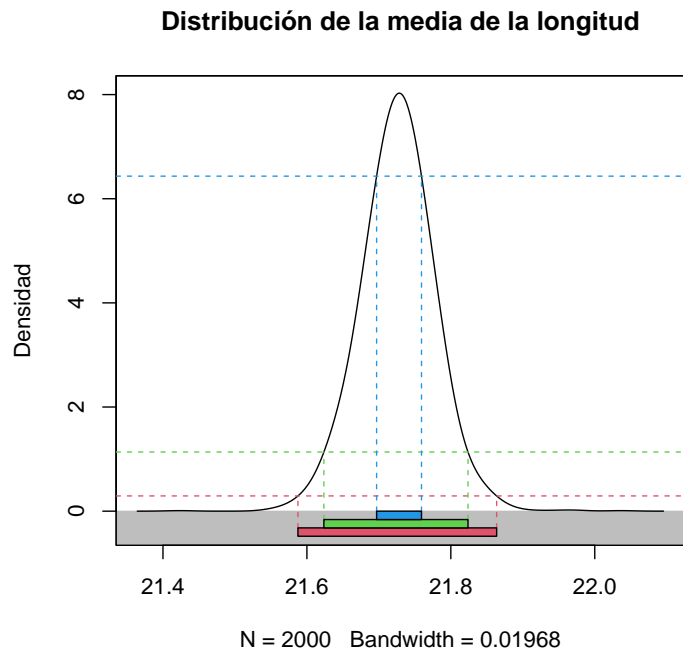


Figura 6.2. *Distribución de la media de la longitud del pie*

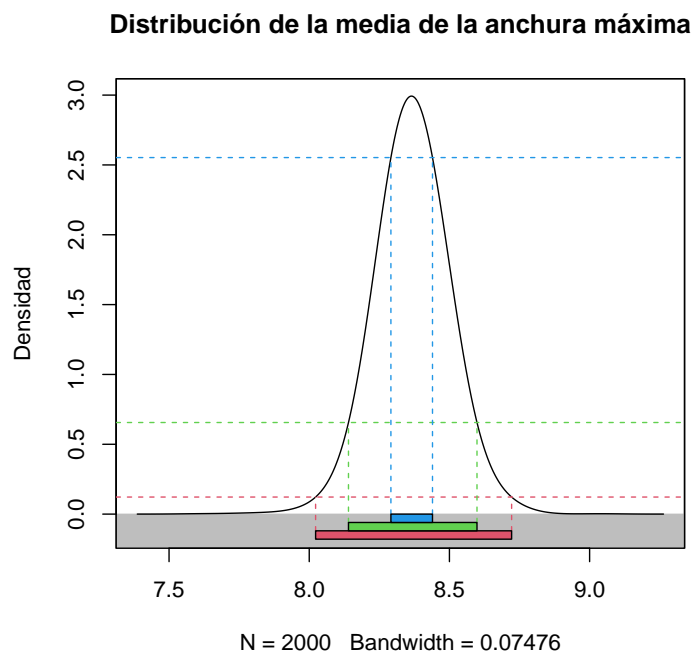


Figura 6.3. *Distribución de la media de la anchura máxima del pie*

```

hdr.den(result[,3],main='Distribución de la media de la anchura talón',ylab='Densidad')
## $hdr
##           [,1]      [,2]
## 99% 4.767364 5.433787
## 95% 4.868756 5.341031
## 50% 5.039910 5.181804

## $mode
## [1] 5.1402

## $falpha
##           1%           5%           50%
## 0.09002895 0.41467171 3.02150251

hdr.den(result[,4],main='Distribución de la media de la longitud del
\ndedo gordo',ylab='Densidad')
## $hdr
##           [,1]      [,2]
## 99% 2.937287 3.625694
## 95% 3.047779 3.521096
## 50% 3.205103 3.351270

## $mode
## [1] 3.27092

## $falpha
##           1%           5%           50%
## 0.0946199 0.4233392 2.7616179

```

Distribución de la media de la anchura talón

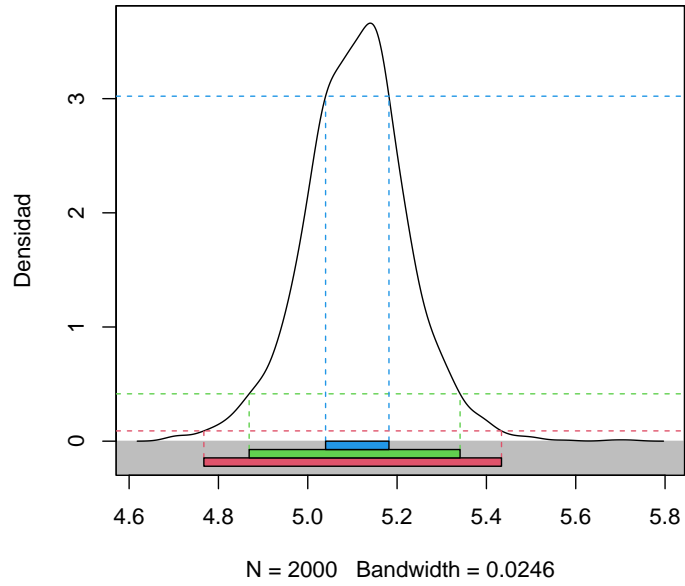


Figura 6.4. *Distribución de la media de la anchura del talón*

**Distribución de la media de la longitud del
dedo gordo**

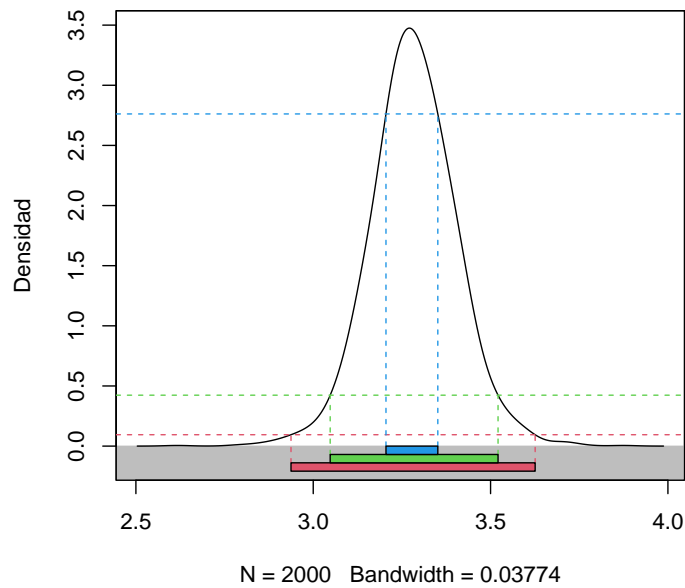


Figura 6.5. *Distribución de la media de la longitud del dedo gordo del pie*

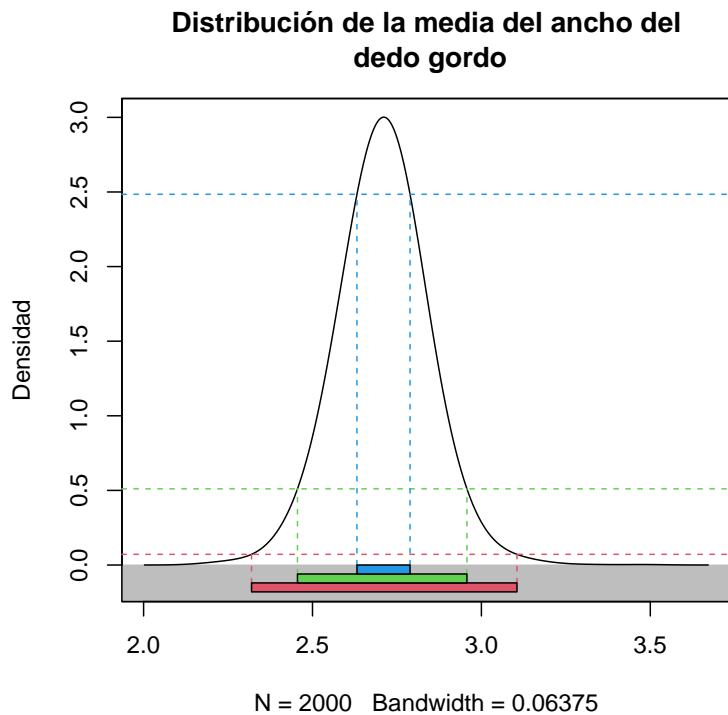


Figura 6.6. *Distribución de la media del ancho del dedo gordo del pie*

```

hdr.den(result[,5],main='Distribución de la media del ancho del\
dedo gordo',ylab='Densidad')
## $hdr
##          [,1]      [,2]
## 99% 2.319520 3.105290
## 95% 2.455494 2.957191
## 50% 2.631605 2.788759

## $mode
## [1] 2.710158

## $falpha
##          1%          5%          50%
## 0.07156966 0.51028359 2.48440481

hdr.boxplot.2d(result[,1],result[,2])
title(main='Distribución conjunta posterior de la longitud\
anchura máximas del pie')
title(ylab='Anchura máxima del pie en cm')
title(xlab='Longitud máxima del pie en cm')

```

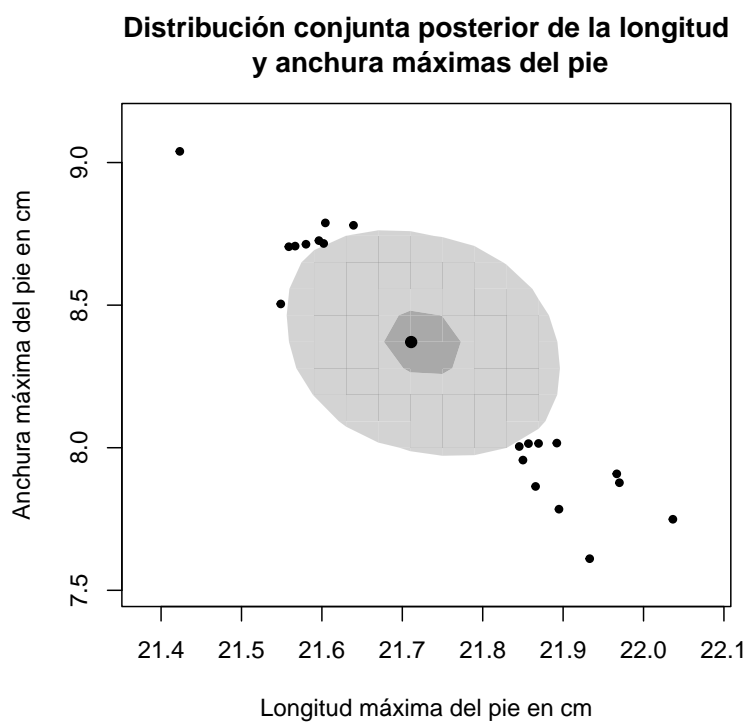


Figura 6.7. *Distribución conjunta posterior de la longitud máxima y anchura máxima del pie*

6.7.6. Elicitación de los parámetros de la a priori conjugada

La elicitación de los parámetros de la distribución conjugada puede ser compleja dada la estructura condicional que se impone sobre esta distribución (Al-Awadhi y Garthwaite, 2001; Garthwaite y Al-Awadhi, 2001), la cual, teóricamente, puede ser muy atractiva, pero desde el punto de vista aplicado en procesos de elicitación no es fácil de manejar, por lo cual ellos proponen el uso de dos distribuciones independientes, una para el vector de medias y otra para la matriz de varianzas y covarianzas. Ellos consideran tres modelos para la a priori.

1. *Modelo 1.* Se especifica un modelo $\xi(\mu, \Sigma)$ tal que:

$$\mu | \Sigma \sim MVN(\mu_0, c^{-1}\Sigma) \quad (6.27)$$

$$\Sigma \sim IW(\delta, A) \quad (6.28)$$

Una matriz aleatoria W se dice que tiene una función densidad de probabilidad llamada *Wishart inversa* si es:

$$f_W(w | \Psi, \nu) = \frac{|\Psi|^{\nu/2}}{2^{\nu/2} \Gamma_p(\nu/2)} |w|^{-(\nu+p+1)/2} \exp\left\{-\frac{1}{2} \text{traza}(\Psi w^{-1})\right\}, \quad (6.29)$$

donde w y Ψ son matrices $p \times p$ definidas positivas, y $\Gamma_p(\cdot)$ es la función gamma multivariable

$$\text{gamma}_p(\cdot) = \pi^{p(p-1)/4} \prod_{j=1}^p \Gamma\left[a + \frac{1-j}{2}\right] \quad (6.30)$$

La media de esta matriz aleatoria es:

$$E[W] = \frac{\Psi}{\nu + p + 1} \quad (6.31)$$

La varianza del ij -ésimo elemento de W es

$$\text{Var}(W_{ij}) = \frac{(\nu - p + 1) \psi_{ij}^2 + (\nu - p - 1) \psi_{ii} \psi_{jj}}{(\nu - p) (\nu - p - 1)^2 (\nu - p - 3)} \quad (6.32)$$

2. *Modelo 2.* Se especifica un modelo $\xi(\mu, \Sigma) = \xi(\mu) \times \xi(\Sigma)$ tal que:

$$\mu \sim MVN(\mu_0, H) \quad (6.33)$$

$$\Sigma \sim IW(\nu, G) \quad (6.34)$$

3. *Modelo 3.* Se especifica un modelo $\xi(\mu, \Sigma) = \xi(\mu) \times \xi(\Sigma)$ tal que:

$$\mu \sim MVN(\mu_0, H) \quad (6.35)$$

$$\Sigma \sim GIW(\nu, G) \quad (6.36)$$

Al-Awadhi y Garthwaite (2001) argumentan que las elicitaciones deben hacerse de tal forma que el experto pueda dar respuestas consistentes con su conocimiento y que realmente correspondan a lo que se quiere realizar. Hacer elicitaciones sobre los parámetros de las distribuciones a priori, directamente, puede llevar a problemas, ya que muchos de los parámetros no tienen en muchas ocasiones interpretaciones claras. Una covarianza, como parámetro, puede ser interpretada a distintos niveles por diferentes expertos, por lo cual podría ser mejor realizar una elicitación sobre un coeficiente de correlación y luego hacer las transformaciones adecuadas. Ellos consideran que las cantidades naturalmente intuitivas son:

- Observaciones directas de la población.
- Promedios de las variables involucradas.
- Desviaciones de observaciones al promedio.

Al-Awadhi y Garthwaite (2001) consideran la *distribución a priori predictiva*, distribución en la cual se eliminan los parámetros vía integración, y que en algunos problemas puede ser más fácil de elicitar que en otros.

Nuestro algoritmo

El algoritmo que proponemos es relativamente sencillo; solo requiere elicitar marginales unidimensionales y distribuciones condicionales entre pares de variables.

1. Para la variable X_i determine los valores elicitados de la media y la varianza, y del nivel de seguridad.
2. La determinación de la covarianza entre X_i y X_j se determina de la siguiente forma:
 - a) Escoja una de las variables como variable explicativa.
 - b) Escoja dos niveles de la variable explicativa.
 - c) Para cada punto de diseño elicite los parámetros de la distribución condicional de la variable dependiente.
 - d) Determine el nivel de seguridad, digamos n_S .
 - e) Genere muestras de la variable dependiente de tamaño n_S para cada nivel de la variable explicativa.
 - f) Estime el modelo lineal.
 - g) De la pendiente determine la covarianza usando la expresión:

$$\beta = \frac{Cov(X_i, X_j)}{\sigma_i}, \quad (6.37)$$

donde σ_i y σ_j se obtienen de las marginales.

Características de los pies de estudiantes. La siguiente tabla presenta las mediciones de sus pies realizadas a 10 estudiantes universitarios.

Tabla 6.4. *Medidas de pies para 10 estudiantes*

x_1	x_2	x_3	x_4	x_5
24.20	9.40	5.50	3.00	3.20
21.70	8.50	6.10	3.20	2.60
25.40	9.60	5.50	4.00	3.10
25.00	10.10	5.30	3.50	2.70
22.00	8.50	5.70	3.10	2.70
25.90	9.30	6.10	4.30	3.30
22.20	8.60	5.20	3.90	2.90
21.70	8.40	5.00	3.20	2.30
25.50	9.20	6.10	3.30	3.20
24.40	9.40	4.70	3.60	2.80

donde:

- x_1 : longitud máxima del pie
- x_2 : amplitud máxima del pie
- x_3 : amplitud máxima del talón
- x_4 : longitud máxima del dedo grande
- x_5 : amplitud máxima del dedo grande

6.8. Distribución a priori potencia

$$\xi^{a_0}(\theta) \propto \xi(\theta) [L(\theta | \text{Datos}_{\text{viejos}})]^{a_0}, \quad (6.38)$$

donde $0 < a_0 < 1$. Este parámetro mide la relevancia de la información muestral disponible para el problema actual.

Si se dispone de una sucesión de datos que fueron obtenidos a través del tiempo, digamos $\text{Datos}_1, \text{Datos}_2, \dots, \text{Datos}_t$, donde los datos fueron obtenidos con modelos parametrizados por θ , entonces:

$$\xi^{a_0}(\theta) \propto \xi(\theta) [L(\theta | \text{Datos}_1)]^{a_1} [L(\theta | \text{Datos}_2)]^{a_2} \times \dots \times [L(\theta | \text{Datos}_t)]^{a_t}, \quad (6.39)$$

donde: $0 < a_1 \leq a_2 \leq \dots \leq a_t < 1$.

6.9. Elicitación de los parámetros del modelo lineal

Modelo lineal bayesiano conjugado

El modelo usual de regresión es:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \tau\mathbf{I}), \quad \boldsymbol{\beta} \in R^k \quad (6.40)$$

La distribución a priori sobre $(\boldsymbol{\beta}, \tau)$

La a priori se expresa como:

$$\xi(\boldsymbol{\beta}, \tau) = \xi(\boldsymbol{\beta}|\tau) \xi(\tau) \quad (6.41)$$

$$\xi(\boldsymbol{\beta}|\tau) = N(\boldsymbol{\beta}_0, \tau\mathbf{T}) \quad (6.42)$$

$$\xi(\tau) = \text{gamma}(\alpha_0, \gamma_0) \quad (6.43)$$

La distribución conjunta ξ de $\boldsymbol{\beta}$ y τ es como sigue para $\boldsymbol{\beta} \in R^k$ y $\tau > 0$:

$$\xi(\boldsymbol{\beta}, \tau) \propto \tau^{k/2} \exp\left(-\frac{\tau}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{T}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)\right) \tau^{\alpha_0-1} \exp(-\gamma_0\tau) \quad (6.44)$$

Los datos y la verosimilitud

Los datos:

$$\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \tau\mathbf{I}) \quad (6.45)$$

Recuerde que $\tau = 1/\sigma^2$.

La función de verosimilitud tendrá la siguiente forma para todo punto $\mathbf{y} = (y_1, \dots, y_n)^T$.

$$L(\boldsymbol{\beta}, \tau | \text{Datos}) \propto \tau^{n/2} \exp\left(-\frac{\tau}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right) \quad (6.46)$$

6.9.1. Ilustración

```
# Ajuste del modelo para el precio de Twingo
# Año (101=2001) y precio (en millones)
datos<-scan()
101 15.2 103 16.9 106 21.4 96 12.3 96 13.0 105 19.9
107 24.5 101 16.5 105 18.9 106 20.5 105 18.7 106 19.0
101 13.8 105 19.0 105 20.0 106 21.5 102 15.5 102 17.5
99 11.0 97 12.5 107 22.5 106 21.5
```

A continuación, presentamos los resultados obtenidos en R ajustando el modelo.

$$\text{Precio de oferta} = \beta_0 + \beta_1 \text{Año} + \epsilon \quad (6.47)$$

```

datos<-matrix(datos ,ncol=2,byrow=T)

Precio<-datos [,2]
Año<-datos [,1]
plot(Año,Precio)

res.lm<-lm(Precio~Año)

summary(res.lm)

## Call:
## lm(formula = Precio ~ Año)

## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8500 -0.8438  0.2416  0.7916  2.8387

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -82.81515     8.82414   -9.385  9.1e-09 ***
## Año           0.97642     0.08559   11.409  3.3e-10 ***
## ---

## Residual standard error: 1.383 on 20 degrees of freedom
## Multiple R-squared:  0.8668,    Adjusted R-squared:  0.8601
## F-statistic: 130.2 on 1 and 20 DF,  p-value: 3.303e-10

```

Construyamos la a priori conjugada. Dependiendo de la cantidad de información encontrada, podemos proceder de diferentes maneras. Asumamos una presentación típica de un modelo estimado que se publica en artículos y reportes. Esto incluye los estimadores de los coeficientes, el error estándar de cada estimador, el valor t para la prueba $H_i : \beta_i = 0$, grados de libertad, $\nu = n - k - 1$, y el MSE, y muchas veces se incluye el coeficiente de determinación.

Si hacemos $Y = \frac{\nu MSE}{\sigma^2} = \nu MSE \tau \sim \chi_\nu^2$. Pero una chi-cuadrado con ν grados de libertad es una gamma con parámetros $\alpha = \frac{\nu}{2}$ y $\beta = 2$. Además, si $W \sim \text{gamma}(\alpha, \beta)$, entonces $kW \sim \text{gamma}(\alpha, k\beta)$. Así, podemos considerar la aproximación:

$$\tau \sim \text{gamma}\left(\frac{\nu}{2}, \frac{2}{\nu MSE}\right) \quad (6.48)$$

La distribución condicional de β se aproxima como:

$$\beta | \tau \sim N\left(\hat{\beta}, \frac{1}{\tau} (X'X)^{-1}\right) \quad (6.49)$$

Dado que la matriz X , generalmente, no aparece en publicaciones, podemos aproximarla así:

1. Para cada variable explicativa determine el rango plausible y cree una rejilla de tamaño n para ella.
2. Permute la rejilla de cada variable en forma independiente y construya una matriz X . Calcule el determinante $|(X'X)^{-1}|$.
3. Conserve la $(X'X)^{-1}$ que maximice el determinante.

6.10. Elicitación de los parámetros del modelo lineal generalizado

Como con el modelo lineal, es posible considerar diferentes procedimientos para hallar la distribución a priori conjunta sobre los parámetros. Hosack et al. (2017) proponen un procedimiento de elicitación que, asumen, es aplicable a los diferentes modelos lineales generalizados. Una aproximación es realizar la elicitación directamente sobre los parámetros, procedimiento que es difícil y que solo es recomendable con personas muy familiarizadas con el significado de los parámetros y sus asociaciones. El método que proponemos aquí es el indirecto, por su facilidad de aplicación y porque el nivel de exigencia al experto es mucho menor. En este caso, se definen unos puntos de diseño (todas las variables explicativas se fijan en valores escogidos por el experto o por la persona que lleva a cabo el proceso) y se genera una elicitación de la distribución de la respuesta en cada punto de diseño.

6.10.1. Regresión logística

El modelo logístico es uno de los más utilizados en trabajos aplicados. Huson y Kinnersley (2009) presentan un método de elicitación para este modelo en caso de estudios de toxicidad, en los cuales a un experto se le pregunta directamente sobre probabilidades condicionales.

Tangarife-Quintero y Correa-Morales (2017) presentan un procedimiento para elicitar los parámetros de un modelo logístico basado en la elicitación de la proporción de éxitos para un conjunto fijo de valores de las variables explicativas (puntos de diseño). Para un modelo logístico con una sola variable explicativa es necesario considerar al menos dos puntos de diseño. Aquí es posible considerar alguna aproximación como las desarrolladas en diseños óptimos. Si tenemos una sola variable explicativa, determinamos de una forma aproximada dónde quedarían los puntos de inflexión del modelo. En estos dos puntos realizamos dos elicitaciones de las distribuciones a priori, como se hace en el caso Bernoulli con una sola proporción. Sin embargo, el experto podría sentirse más cómodo en valores distintos de la variable explicativa.

Algoritmo para un modelo logístico con una sola variable explicativa

1. Fije al menos dos puntos de la variable explicativa en los que el experto se sienta cómodo y no se tengan probabilidades de éxito extremas (muy cerca de uno o de cero).
2. Para un punto x , determine la distribución a priori para el parámetro π_x . Por comodidad, se puede aproximar la a priori mediante una distribución *beta* (α_x, β_x) y un nivel de seguridad n_x .
3. Para cada distribución a priori se genera una proporción; digamos que para el punto x generamos $\pi_x^{(l)}$. Luego sacamos una muestra de la distribución *Bernoulli* ($\pi_x^{(l)}$) de tamaño n_x y tenemos la muestra $(x, y_i^x)_{i=1}^{n_x}$.
4. Con las muestras de los puntos de la variable explicativa seleccionados ajustamos el modelo logístico.

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 \cdot x, \quad (6.50)$$

y guardamos (β_0^l, β_1^l) , $l = 1, 2, \dots, Nsim$.

5. Con estos vectores de parámetros estimados calculamos el vector de medias y la matriz de covarianzas de los parámetros.

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \sim Normal\left(\mu_\beta = \begin{pmatrix} \mu_{\beta_0} \\ \mu_{\beta_1} \end{pmatrix}; \Sigma_\beta = \begin{pmatrix} \sigma_{\beta_0}^2 & \sigma_{\beta_0, \beta_1} \\ \sigma_{\beta_0, \beta_1} & \sigma_{\beta_1}^2 \end{pmatrix}\right) \quad (6.51)$$

6.10.2. Regresión de Poisson

Procediendo de forma similar al caso del modelo logístico, la elicitación de los parámetros en un modelo de regresión de Poisson se basa en la determinación de valores esperados de la variable respuesta para diferentes niveles de la variable explicativa. Para definir un modelo de esta naturaleza con una única covariable, es necesario seleccionar al menos dos puntos de diseño adecuados. En este contexto, es posible emplear aproximaciones inspiradas en diseños óptimos para optimizar la selección de estos puntos.

Dado que el modelo de Poisson describe conteos, se busca identificar aquellos valores de la variable explicativa donde la variabilidad en el conteo de eventos sea mayor. En esos puntos, se elicitán distribuciones a priori para el parámetro de la media del conteo, siguiendo un procedimiento análogo al empleado en la elicitación de proporciones en modelos Bernoulli. Sin embargo, el experto podría preferir puntos distintos si ello facilita una mejor interpretación del fenómeno bajo estudio.

Procedimiento para un modelo de Poisson con una variable explicativa

1. Seleccionar al menos dos valores de la variable explicativa que sean adecuados según el criterio del experto. Estos deben ubicarse en un intervalo razonable, evitando valores en los que la ocurrencia de eventos sea extremadamente improbable.
2. Para cada punto x , establecer una distribución a priori para el parámetro de la media del conteo, denotado como λ_x . Una elección conveniente es la distribución $gamma(\alpha_x, \beta_x)$, junto con un nivel de confianza n_x .
3. A partir de la distribución a priori seleccionada, generar valores simulados de λ_x , es decir, obtener $\lambda_x^{(l)}$ para cada iteración. Posteriormente, extraer una muestra de tamaño n_x a partir de una distribución de Poisson con media $\lambda_x^{(l)}$, obteniendo la muestra $(x, y_i^x)_{i=1}^{n_x}$.
4. Utilizando las muestras generadas para los diferentes valores de la variable explicativa, ajustar el modelo de regresión de Poisson:

$$\log(\lambda) = \beta_0 + \beta_1 x, \quad (6.52)$$

y registrar los pares de parámetros estimados (β_0^l, β_1^l) para cada iteración $l = 1, 2, \dots, Nsim$.

5. Con las estimaciones obtenidas, calcular el vector de medias y la matriz de covarianzas de los parámetros:

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \sim normal \left(\mu_\beta = \begin{pmatrix} \mu_{\beta_0} \\ \mu_{\beta_1} \end{pmatrix}, \Sigma_\beta = \begin{pmatrix} \sigma_{\beta_0}^2 & \sigma_{\beta_0, \beta_1} \\ \sigma_{\beta_0, \beta_1} & \sigma_{\beta_1}^2 \end{pmatrix} \right) \quad (6.53)$$

Ejemplo. El número de veces que un estudiante masculino usó los servicios sanitarios el día anterior puede dar una idea del número de baños que se necesitan en el campus. Si asumimos que el modelo de Poisson puede usarse para estudiar este conteo, considerando variable explicativa el número de horas que el estudiante estuvo presente en el campus, entonces determinamos la a priori que nos resume la información de un experto. Suponga que escogemos los dos puntos: $x = 4$ (horas) y $x = 8$. En estos puntos realizamos la elicitación.

- **Pregunta 1.** De 1000 estudiantes que están exactamente cuatro horas en el campus, ¿cuántos cree usted no van al baño?
- **Pregunta 2.** De esos 1000, ¿cuántos van al baño solo una vez?
- **Pregunta 3.** ¿Cuántos van dos veces?
- **Pregunta 4.** ¿Cuántos van tres veces?

- **Pregunta 5.** ¿Cuántos van cuatro?
- **Pregunta 6.** ¿Cuántos van cinco?
- etc.

Suponga que el experto nos genera esta tabla:

Idas al baño	0	1	2	3	4	5	6	7	8
Número de estudiantes	400	350	150	50	35	10	4	1	0

Determinamos el nivel de seguridad del experto con relación a la tabla anterior, digamos $n_4 = 30$.

Obtenemos muchas muestras de tamaño $n_4 = 30$, digamos $Nsim = 1000$ y para cada una obtenemos el promedio, digamos $\{\bar{x}_i\}_{i=1}^{Nsim}$.

A estas medias les ajustamos una distribución $gamma(\alpha_4, \beta_4)$.

Repetimos el paso anterior con los otros niveles escogidos para la variable explicativa. Terminamos con una tabla como la siguiente:

Nivel de covariable	α	β	Nivel de seguridad
$x = 4$	$\alpha_4 = 23.270277$	$\beta_4 = 22.476849$	$n_4 = 30$
$x = 8$	$\alpha_8 = 30.3517201$	$\beta_8 = 14.5711617$	$n_8 = 10$

Sacamos un número grande de muestras simuladas de la distribución $gamma$ para cada nivel de la covariable (igual número para cada nivel). Cada uno de estos valores corresponde a la media de una Poisson de la cual generamos una muestra de tamaño igual al nivel de seguridad de ese nivel. Con estas muestras para cada nivel formamos.

$$\begin{array}{ll}
 x_1 & y_{1,1} \\
 x_1 & y_{2,1} \\
 \vdots & \vdots \\
 x_1 & y_{n_1,1} \\
 x_2 & y_{1,2} \\
 x_2 & y_{2,2} \\
 \vdots & \vdots \\
 x_2 & y_{n_2,2}
 \end{array}$$

Con estos datos ajustamos el modelo de Poisson y guardamos (β_0, β_1) .

```

# Niveles
x1<-4
x2<-8
idas.al.baño<-0:8
frecu.x1<-c(400,350,150,50,35,10,4,1, 0)
nivel.segu1<-30

frecu.x2<-c(50,250,450,150,50,35,10,5, 0)
nivel.segu2<-10

Nsim1<-1000
medias1<-apply(matrix(sample(idas.al.baño,Nsim1*nivel.segu1,
replace=T,prob=frecu.x1),ncol=nivel.segu1),1,mean)
medias2<-apply(matrix(sample(idas.al.baño,Nsim1*nivel.segu2,
replace=T,prob=frecu.x2),ncol=nivel.segu2),1,mean)

library(MASS)
fitdistr(medias1,'gamma')
fitdistr(medias2,'gamma')

Nsim2<-1000
medias11<-rgamma(Nsim2,23.270277,rate=22.476949)
medias22<-rgamma(Nsim2,30.351702,rate=14.5711617)

temporal<-function(medias,n.s){
y<-c(rpois(n.s[1],medias[1]),rpois(n.s[2],medias[2]))
x<-c(rep(4,n.s[1]),rep(8,n.s[2]))
resu<-glm(y~x,family='poisson')$coefficients
return(resu)
}

resu<-apply(cbind(medias11,medias22),1,temporal,c(nivel.segu1,nivel.
segu2))
dim(resu)
colMeans(t(resu))
cov(t(resu))

colMeans(t(resu))
## (Intercept)          x
## -0.7243575    0.1779157
cov(t(resu))
##          (Intercept)          x
## (Intercept)  0.42093325 -0.06377477
## x          -0.06377477  0.01072094

```

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \sim Normal \left(\mu_\beta = \begin{pmatrix} -0.7243575 \\ 0.1779157 \end{pmatrix}; \Sigma_\beta = \begin{pmatrix} 0.42093325 & -0.06377477 \\ -0.06377477 & 0.01072094 \end{pmatrix} \right) \quad (6.54)$$

Elicitación en series de tiempo: Durbach et al. (2017) proponen un protocolo para la elicitación de pronósticos a largo plazo.

6.10.3. Conclusiones y recomendaciones

A pesar de contar con un solo estudio como información previa, los métodos de elicitación presentados en este capítulo permiten hacer estimaciones confiables de los parámetros de interés. De cualquier modo, es recomendable establecer siempre métodos de validación de la robustez de las estimaciones obtenidas con el fin de evaluar la distribución a priori elicitada, y seguir de manera cuidadosa los pasos o procedimientos necesarios para elicitar distribuciones a priori.

Capítulo 7

Elicitación a partir de varios estudios: metaanálisis

Cuando existen estudios previos cuyos resultados han sido publicados, usualmente uno no tiene acceso a los datos primarios, sino a resúmenes más o menos completos, por ejemplo, estimadores, intervalos de confianza, etc. Por ejemplo, en un estudio previo sobre la estatura promedio de hombres adultos jóvenes a partir de una muestra de 543 jugadores del mundial de Suráfrica se obtuvo un valor de 181.3131 y una desviación típica de 6.568231; por lo tanto, la distribución a priori para la estatura media podría ser una $N(181.3131, 43.14166/543)$. Si los datos fueran muy antiguos, se podría considerar una penalización que aumente la varianza por una constante que dependa de lo antiguos que sean ellos, por ejemplo, $N(181.3131, \kappa_t 43.14166/543)$, donde $\kappa_t > 1$ es el parámetro de penalización.

Si se tiene acceso a los datos primarios del estudio previo, se puede realizar un análisis bayesiano tradicional considerando, posiblemente, distribuciones a priori no informativas.

Elementos que se deben considerar cuando se utilizan estudios previos:

- *Pertinencia del estudio.* Debemos considerar estudios que hayan apuntado al mismo conjunto de parámetros que son de nuestro interés y que se hayan realizado dentro de poblaciones similares a la que vamos a considerar.
- *Oportunidad del estudio.* El periodo durante el cual el estudio fue realizado; este debería estar lo suficientemente cerca al nuestro, de tal forma que garanticemos que no ha habido un cambio estructural en el problema que estamos manejando. Por ejemplo, suponga que estamos interesados en estimar la proporción de mujeres adultas que escuchan radionovelas en la actualidad y, suponga, que conseguimos un estudio realizado antes de la aparición de la televisión en los años 50.

Consideremos el estudio realizado por Barrientos et al. (2014) acerca de la prevalencia de insomnio y somnolencia entre estudiantes de Medicina en una institución universitaria. El estudio, basado en 230 estudiantes, presenta los siguientes datos sobre somnolencia:

Tabla 7.1. *Distribución porcentual de la prevalencia de insomnio y somnolencia en 230 estudiantes*

Estado	Porcentaje
Sin problema	26.5
Leve	17.0
Moderado	46.1
Grave	10.4

Utilizando los datos proporcionados, generamos M muestras de tamaño 230 de una multinomial con proporciones 0.265, 0.170, 0.461, 0.104. Con estas muestras estimamos los parámetros de una a priori Dirichlet.

7.1. Metaanálisis

En muchas situaciones, los investigadores, en la búsqueda de antecedentes para resolver una cuestión de investigación, se encuentran con estudios realizados y que apuntan al mismo asunto. Los investigadores pueden estar interesados en combinar los resultados de estos estudios, experimentales u observacionales, de tal forma que se obtenga un solo resultado al final. Metaanálisis hace referencia al proceso de localizar, seleccionar, asignar y combinar información relevante a un asunto de investigación particular. Esta es una visión amplia del metaanálisis, ya que algunos investigadores lo entienden solo como el proceso de resumen, por ejemplo Villar, J. et al. (2001): “Entendemos el metaanálisis como el uso de técnicas estadísticas para combinar los resultados de estudios que miran la misma cuestión en una medida de resumen”.

El metaanálisis es una metodología muy controvertida entre investigadores. Villar, J. et al. (2001) dicen que revisiones sistemáticas, incluyendo el metaanálisis, han tenido un auge en su uso durante los últimos años en la literatura médica. Hay una necesidad de realizar más trabajo empírico metodológico para entender las propiedades y limitaciones de las técnicas de la metodología estadística.

Si el número de estudios independientes que han sido llevados a cabo es un asunto particular, utilizando datos y métodos diferentes, entonces combinando sus resultados puede proveer más entendimiento y mayor poder explicatorio que simplemente listar los resultados individuales. En verdad, el metaanálisis ha llegado a ser una práctica aceptada para evaluar el actual flujo de evidencia científica contradictoria.

Descubrir significancia estadística en un área controversial que está plagada con diversos y ambiguos hallazgos, ha sido el objetivo muchos metaanálisis en las ciencias sociales e investigación médica (Stanley, 2001).

¿Cómo debe un revisor conducir un metaanálisis? ¿Qué problemas o consideraciones deben tenerse en cuenta? Un paso fundamental es tener una lista y definición clara de cada uno de los parámetros a elicitar.

- Paso 1: incluya todos los estudios relevantes de una base de datos estándar.

El metaanálisis comienza con el énfasis absoluto de incluir todos los estudios, publicados o no, como una forma de reducir sesgos potenciales introducidos por una selección no aleatoria de estudios. Del Re (2015) recomienda tener criterios absolutamente claros de inclusión/exclusión basados en las hipótesis iniciales que, en nuestro caso, serían las definiciones de los parámetros de interés. (Presunción de inclusividad).

- Paso 2: seleccione un estadístico de resumen y reduzca la evidencia a una métrica común.

Para resumir la evidencia de diferentes pruebas estadísticas, es necesario transformar los estadísticos de resumen a una métrica común y comparable. Las pruebas estadísticas usadas (razón de verosimilitud, multiplicador de Lagrange, o Wald) varían de estudio a estudio y se distribuyen como chi-cuadrado o F. Sin embargo, ambas distribuciones dependen de los grados de libertad disponibles, que difieren entre estudios y, por lo tanto, no pueden ser comparados, resumidos o analizados directamente; así, alguna transformación a una métrica común se hace necesaria. La conversión a la normal estándar es la forma más obvia y práctica para tratar con estas incompatibilidades.

A veces no es claro cuando se usa la palabra metaanálisis si corresponde a una síntesis cuantitativa o a la combinación de una revisión sistemática de la literatura y de síntesis cuantitativa (Petitti, 2001). El entusiasmo inicial por el metaanálisis era debido, en parte, a la creencia, de que teniendo la información pertinente, sería posible obtener un solo valor estimado de resumen del efecto de las intervenciones o de la relación entre factores de riesgo y enfermedad.

La idea básica del metaanálisis es combinar información provenientes de diferentes estudios pero realizados esencialmente sobre el mismo fenómeno (ensayos clínicos, evaluaciones económicas, estudios epidemiológicos, etc.), con el propósito de tener unas inferencias y predicciones más exactas que las que se tengan de cualquier estudio individual. Aquí tenemos sujetos dentro de estudios y habrá predictores a la vez de los sujetos y de los estudios.

Los objetivos del metaanálisis pueden ser:

1. Establecer protocolos sistemáticos para el diseño de estudios y el registro de resultados,
2. Combinar la evidencia de diferentes estudios que individualmente pueden estar sujetos a variabilidad muestral.

El énfasis se hace en combinar estimadas del tamaño de los efectos de un número de estudios, tales efectos pueden ser razones de odds, diferencias de riesgos, razones de riesgos, etc. También puede ser más complejo, tal como combinar regresiones, coeficientes de correlación, etc.

Villar, J. et al. (2001) comentan que aunque hay una tendencia en la literatura a forzar los resultados de una revisión sistemática en un solo estimador combinado usando técnicas metaanálisis, es claro que cuando existen diferencias sustanciales entre los resultados de los ensayos, y enfrentamos heterogeneidad estadística, un solo valor del estimación puede darnos una impresión incorrecta y esto debe evitarse.

Un metaanálisis de ensayos clínicos debe reportar, en general, más que un solo valor del estimador mezclado del efecto del tratamiento. Si existe heterogeneidad entre los diferentes ensayos, como a menudo sucede, debe hacerse el intento de explorar sus posibles causas. La mejor forma sería usar los datos de los pacientes individuales para explorar el efecto de covariables particulares sobre el efecto del tratamiento. Sin embargo, es común en metaanálisis no tener acceso a estos datos individuales y solo se tienen los resúmenes, tales como medias, frecuencias y proporciones (Sharp, S. J. y Thompson, S. G., 2000).

Ejemplo 7.1. *Stangl (2001) presenta el siguiente caso en el que se realizó un análisis de los efectos de la droga antidepresiva S-adenosylometionina (SAMe). Participaron nueve lugares en el ensayo. Cada sitio tenía características propias que afectaban los resultados de los estudios. El resultado de interés era la tasa de éxito observada con la SAMe. Los datos están presentados en la siguiente tabla:*

Tabla 7.2. *Tasa de éxito observada con la SAMe para nueve lugares de ensayo*

Sitio	s_i	n_i	s_i/n_i
1	20	20	1.00
2	4	10	0.40
3	11	16	0.69
4	10	19	0.53
5	5	14	0.36
6	36	46	0.78
7	9	10	0.90
8	7	9	0.78
9	4	6	0.67
Total	106	150	0.71

Asumimos que la tabla anterior proviene de un estudio con 150 pacientes y nuestro objetivo es estimar la tasa de éxito, digamos π , del tratamiento. Asumamos que la distribución a priori de π es una *beta* (α, β). Los datos son generados de una distribución binomial con tamaño muestral n y tasa de éxito π .

Una formulación multinivel, jerárquica o de efectos aleatorios evita el supuesto de homogeneidad modelando un efecto aleatorio π para el estudio i . Cada π_i se asume sacado de la distribución de los efectos de estudio. Aquí se usa la *beta* (α, β) para los efectos de estudio. La respuesta del estudio i es

$$s_i \sim \text{Binomial}(n_i, \pi_i), \quad (7.1)$$

y los efectos individuales son intercambiables. Condicionados en α y β , los π_i son sacados independientemente de una distribución *beta*:

$$\pi_i \sim \text{beta}(\alpha, \beta) \quad (7.2)$$

La función de verosimilitud de los π_i 's es:

$$\prod_{i=1}^I \pi_i^{s_i} (1 - \pi_i)^{n_i - s_i} \quad (7.3)$$

En el último nivel de jerarquía, una distribución a priori es ubicada sobre α y β . La dejamos sin especificar por el momento, y simplemente la denotamos por:

$$\xi(\alpha, \beta) \quad (7.4)$$

La distribución posterior conjunta de todos los parámetros es:

$$\begin{aligned} \xi(\pi, \alpha, \beta | s) &\propto f(s | \pi, \alpha, \beta) \xi(\pi | \alpha, \beta) \xi(\alpha, \beta) \\ &\propto \prod_{i=1}^I \pi_i^{s_i} (1 - \pi_i)^{n_i - s_i} \prod_{i=1}^I \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \pi_i^{\alpha-1} (1 - \pi_i)^{\beta-1} \xi(\alpha, \beta) \end{aligned} \quad (7.5)$$

Dados α y β , cada uno de los π_i tiene una distribución beta independiente. Su densidad conjunta será:

$$\xi(\pi | \alpha, \beta, s) \propto \prod_{i=1}^I \frac{\Gamma(\alpha + \beta + n_i)}{\Gamma(\alpha + s_i)\Gamma(\beta + n_i - s_i)} \pi_i^{\alpha + s_i - 1} (1 - \pi_i)^{\beta + n_i - s_i - 1} \quad (7.6)$$

La marginal posterior de (α, β) es:

$$\xi(\alpha, \beta | s) \propto \xi(\alpha, \beta) \prod_{i=1}^I \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + s_i)\Gamma(\beta + n_i - s_i)}{\Gamma(\alpha + \beta + n_i)} \quad (7.7)$$

Ejemplo 7.2. *Bravo et al. (2003)* realizaron una revisión de los estudios sobre la presencia del *Helicobacter pylori*, la bacteria asociada con gastritis.

Tabla 7.3. Prevalencia del *Helicobacter pylori* en diferentes regiones de Colombia (*Bravo et al., 2003*)

	Año	Positivos	Total	Proporción
1	87	173	244	0.71
2	88	109	135	0.81
3	89	110	164	0.67
4	89	66	99	0.67
5	89	47	68	0.69
6	89	53	62	0.85
7	89	49	53	0.92
8	89	22	35	0.63
9	90	75	90	0.83
10	90	54	54	1.00
11	90	62	64	0.97
12	91	1	30	0.03
13	91	18	60	0.30
14	91		30	
15	92	24	50	0.48
16	94	63	90	0.70
17	94	10	18	0.56
18	96	60	60	1.00
19	96	22	50	0.44
20	96	30	49	0.61
21	96	855	878	0.97
22	96	472	684	0.69

```
temp<-scan()
1 87 173 244 0.71    2 88 109 135 0.81
3 89 110 164 0.67    4 89 66 99 0.67
5 89 47 68 0.69     6 89 53 62 0.85
7 89 49 53 0.92     8 89 22 35 0.63
9 90 75 90 0.83    10 90 54 54 1.00
11 90 62 64 0.97   12 91 1 30 0.03
13 91 18 60 0.30   15 92 24 50 0.48
16 94 63 90 0.70   17 94 10 18 0.56
18 96 60 60 1.00   19 96 22 50 0.44
20 96 30 49 0.61   21 96 855 878 0.97
22 96 472 684 0.69
```

```

temp<-matrix(temp,ncol=5,byrow=T)
eventos<-temp[,3]
ns<-temp[,4]

library(meta)
res<-metaprop(eventos,ns)
summary(res)
## Number of studies combined: k = 21

##              proportion          95%-CI
## Fixed effect model      0.7820 [0.7670; 0.7963]
## Random effects model    0.7694 [0.6111; 0.8763]

## Quantifying heterogeneity:
## tau^2 = 2.8792; tau = 1.6968; I^2 = 98.1%; H = 7.21

## Test of heterogeneity:
##      Q d.f.  p-value          Test
## 298.62  20 < 0.0001      Wald-type
## 635.69  20 < 0.0001 Likelihood-Ratio

## Details on meta-analytical method:
## - Random intercept logistic regression model
## - Maximum-likelihood estimator for tau^2
## - Logit transformation
## - Continuity correction of 0.5 in studies with zero cell
  frequencies (only used to calculate individual study results)

```

Los resultados obtenidos utilizando el modelo mixto son preferidos en términos generales, ya que se considera que los que aparecen en la literatura provienen de experimentos realizados en situaciones diferentes y por investigadores distintos (Del Re, 2015). Un caso en el que se podría considerar un modelo de efectos fijos es ilustrado con un ejemplo trivial: suponga que una moneda es lanzada por k sujetos un número de veces distintas y deseamos estimar la probabilidad de obtener cara. Volviendo a nuestro problema, si deseamos resumir la información disponible obtenida por el metaanálisis y asumiendo que los estudios tienen una extravariabilidad, debemos hallar los valores de los parámetros de la distribución escogida por nosotros, por ejemplo la $beta(\alpha_0, \beta_0)$. Podemos proceder así:

$$\frac{\alpha_0 - 1}{\alpha_0 + \beta_0 - 2} = 0.7694$$

$$F(0.6111 | \alpha_0, \beta_0) = 0.025$$

$$F(0.8763 | \alpha_0, \beta_0) = 0.975,$$

donde $F(\cdot)$ es la función de distribución acumulada de una $beta(\alpha_0, \beta_0)$. Hallar los valores (α_0, β_0) se hace numéricamente.

```

a.minimizar<-function(parame){
  a0<-parame[1]
  b0<-parame[2]
  temp<-(0.7694-(a0-1)/(a0+b0-2))^2+(0.6111-qbeta(0.025,a0,b0))^2
    +(0.8763-qbeta(0.975,a0,b0))^2
  return(temp)
}

optim(c(2,2),a.minimizar)
## $par
## [1] 29.215898 9.435129

## $value
## [1] 4.289818e-07

## $counts
## function gradient
##      67      NA

## $convergence
## [1] 0

## $message
## NULL

```

Por lo tanto, los estudios previos generan una a priori $\beta(29.215898; 9.435129)$. La forma de elicitación es directa y simple.

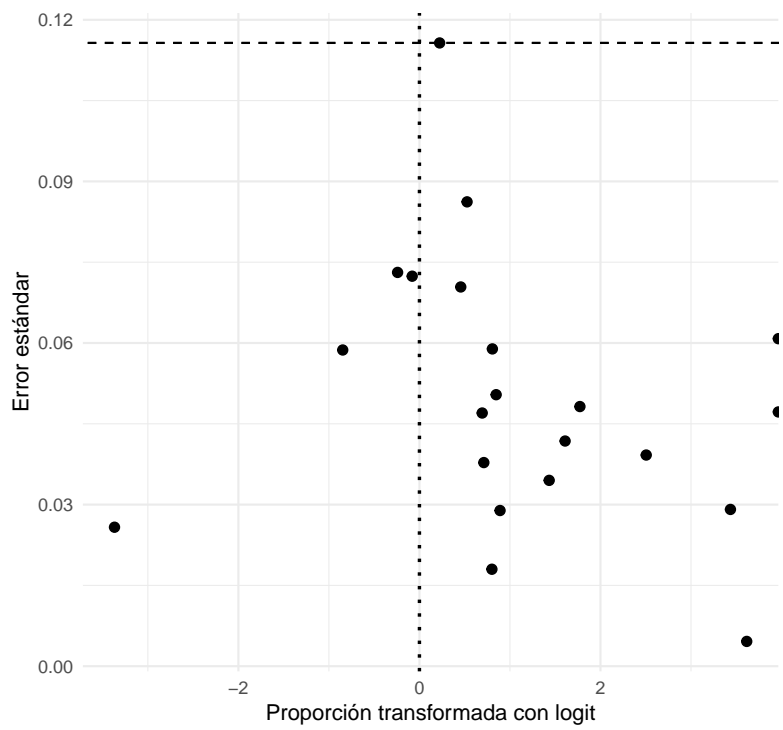


Figura 7.1. Errores estándar de la proporción estimada (transformada) para la presencia del *Helicobacter pylori*

En el caso de la variabilidad se verifica usando el estadístico Q definido como:

$$Q = \sum_{i=1}^k w_i (y_i - \hat{\theta})^2 \sim \chi_{k-1}^2, \quad (7.8)$$

donde k es el número de estudios considerados y w_i es la ponderación del i -ésimo estudio. Otras medidas de variabilidad son:

$$I^2 = \left(\frac{Q - gl}{Q} \right) 100 \% \quad (7.9)$$

7.1.1. Análisis estadístico

Peters, U., Poole, C. y Arab, L. (2001) indican que los análisis estadísticos incluyen:

1. Extraer o calcular una estimación común y comparable del riesgo relativo de cada estudio.
2. Búsqueda de evidencia de sesgo de publicación.
3. Análisis de la variación interestudios.
4. Cálculo de de las estimaciones de resumen del efecto. La evidencia de sesgo de publicación y la heterogeneidad debe tomarse como un contraindicativo de la confianza en los resúmenes globales. En presencia de heterogeneidad del efecto, el principal propósito del metaanálisis será la identificación de las fuentes de variación entre estudios.

Los métodos estadísticos están generalmente basados en modelos de efectos fijos o de efectos aleatorios.

7.1.2. Modelo de efectos fijos

Supongamos que tenemos una colección de k estudios, donde el i -ésimo tiene un efecto estimado de tamaño Y_i y el verdadero tamaño del efecto es θ_i . Un modelo general está dado por:

$$Y_i = \theta_i + e_i, \quad (7.10)$$

donde $e_i \sim N(0, \sigma_i^2)$, $i = 1, 2, \dots, k$. Estos componentes son desviaciones aleatorias del verdadero tamaño del efecto y se asume que son independientes. Esto implica que $Y_i \sim N(\theta_i, \sigma_i^2)$, $i = 1, 2, \dots, k$. Aunque el modelo anterior asume normalidad, se aplica también a situaciones en las que se tiene normalidad aproximada, por ejemplo, cuando se utiliza la razón de odds o riesgos relativos.

En general, el parámetro de interés es el efecto global, denotado por μ . El modelo de efectos fijos asume que $\theta_i = \mu$ para todo $i = 1, 2, \dots, k$, implicando que cada

estudio en el metaanálisis tiene el mismo efecto base. Note que, aun teniendo el mismo efecto base, las Y_i no se distribuyen idénticamente, ya que tienen diferentes varianzas. El estimador de μ es generalmente un promedio ponderado simple de las Y_i , con ponderaciones óptimas proporcionales a $w_i = 1/\text{var}(Y_i)$. En la práctica, las varianzas son desconocidas y se utilizan las estimaciones de ellas, digamos $\hat{\sigma}_i^2$, las cuales son usadas para estimar a μ y a $\text{var}(\hat{\mu})$. Tenemos entonces:

$$\hat{\mu} = \frac{\sum_{i=1}^k \hat{w}_i Y_i}{\sum_{i=1}^k \hat{w}_i}, \quad (7.11)$$

donde $\hat{w}_i = 1/\hat{\sigma}_i^2$, y además:

$$\hat{\text{var}}(\hat{\mu}) = \frac{1}{\sum_{i=1}^k \frac{1}{\hat{\sigma}_i^2}} \quad (7.12)$$

El modelo de efectos mixtos no asume que las θ_i son iguales, sino que ellas se distribuyen normalmente. Esto nos lleva a un modelo de dos etapas:

$$Y_i = \theta_i + e_i \quad (7.13)$$

$$\theta_i = \mu + \epsilon_i, \quad (7.14)$$

donde $e_i \sim N(0, \sigma_i^2)$ y $\epsilon_i \sim N(0, \tau^2)$, $i = 1, 2, \dots, k$. Además, se asume que las e_i y las ϵ_i son independientes. En este caso, el verdadero efecto para el estudio i está centrado alrededor del efecto global, permitiendo que los estudios individuales varíen tanto en el estimado como en el verdadero efecto. El parámetro τ^2 , la varianza de los efectos aleatorios, es una medida de la heterogeneidad entre estudios.

7.1.3. Modelo de efectos aleatorios

El modelo de efectos aleatorios puede escribirse como:

$$Y_i = \mu + \epsilon_i + e_i, \quad (7.15)$$

donde $e_i \sim N(0, \sigma_i^2)$ y $\epsilon_i \sim N(0, \tau^2)$, $i = 1, 2, \dots, k$, que relaciona las Y_i directamente a la medida global del efecto μ , dado por:

$$\hat{\mu}_\tau = \frac{\sum_{i=1}^k \hat{w}_i(\tau) Y_i}{\sum_{i=1}^k \hat{w}_i(\tau)}, \quad (7.16)$$

con varianza

$$\text{var}(\hat{\mu}_\tau) = \frac{1}{\sum_{i=1}^k \hat{w}_i(\tau)}, \quad (7.17)$$

donde $\hat{w}_i(\tau) = (\tau^2 + \hat{w}_i^{-1})^{-1}$ y \hat{w}_i definida antes. Asumiendo que τ^2 sea conocida, tenemos entonces:

$$\hat{\mu}_\tau \sim N \left(\mu, \frac{1}{\sum_{i=1}^k \hat{w}_i(\tau)} \right) \quad (7.18)$$

En la práctica τ^2 es desconocido. El estimador más popular es de DerSimonian y Laird, el cual se deriva a partir de la idea de igualar el valor esperado de $Q_{\hat{w}}$ con su valor observado. Note que:

$$E(Q_{\hat{w}}) = k - 1 + \tau^2 \left(\sum_{i=1}^k \hat{w}_i - \frac{\sum_{i=1}^k \hat{w}_i^2}{\sum_{i=1}^k \hat{w}_i} \right) \quad (7.19)$$

Suponga que t se obtiene resolviendo:

$$q_{\hat{w}} = k - 1 + t \left(\sum_{i=1}^k \hat{w}_i - \frac{\sum_{i=1}^k \hat{w}_i^2}{\sum_{i=1}^k \hat{w}_i} \right), \quad (7.20)$$

lo cual produce:

$$t = \frac{q_{\hat{w}} - (k - 1)}{\sum_{i=1}^k \hat{w}_i - \frac{\sum_{i=1}^k \hat{w}_i^2}{\sum_{i=1}^k \hat{w}_i}} \quad (7.21)$$

Es posible que $t < 0$, en este caso, se toma igual a cero. Esta restricción produce un estimado sesgado para τ^2 .

7.1.4. Heterogeneidad

Villar, J. et al. (2001) comentan que aunque hay una tendencia en la literatura a forzar los resultados de una revisión sistemática en un solo estimador combinado usando técnicas del metaanálisis, es claro que cuando existen diferencias sustanciales entre los resultados de los ensayos, y enfrentamos heterogeneidad estadística, un solo valor de la estimación puede darnos una impresión incorrecta y esto debe evitarse.

Según Sharp, S. J. y Thompson, S. G. (2000) un metaanálisis de ensayos clínicos debe reportar en general más que un solo valor del estimador mezclado del efecto del tratamiento. Si existe heterogeneidad entre los diferentes ensayos, como a menudo sucede, debe hacerse el intento de explorar sus posibles causas. La mejor forma sería usar los datos de los pacientes individuales para explorar el efecto de covariables particulares sobre el efecto del tratamiento. Sin embargo, es común en metaanálisis no tener acceso a estos datos individuales y solo se tienen los resúmenes, tales como medias, frecuencias y proporciones.

Prueba Q de Cochran. La heterogeneidad es normalmente verificada utilizando la prueba de Cochran cuyo estadístico de prueba es:

$$Q_w = \sum_{i=1}^k w_i (Y_i - \hat{\mu})^2 \quad (7.22)$$

Con H_0 se tiene que $Q_w \sim \chi_{k-1}^2$. En la práctica se trabaja con \hat{w}_i ; por lo tanto, la prueba será aproximada.

7.1.5. Sesgo de publicación

Stanley, T. D. (2001) plantea que es común en la revisión de literatura que el autor descalifique ciertos estudios como “mal especificados”, un juicio que muy a menudo parece hacerse de una forma subjetiva y selectiva ya que se descalifican aquellos que van en contra de las creencias del autor. Rust, Lehmann y Farley (1990) comentan que a medida que un campo de investigación madura, el “sesgo de publicación” puede incrementarse, esto debido a la tendencia de las revistas a aceptar solo hallazgos estadísticamente significativos o con fuertes efectos, lo que puede originar un sesgo en la magnitud de los efectos reportados. También los autores, condicionados por los revisores, pueden suprimir o minimizar los resultados que muestren efectos pequeños o no significativos.

Eberly, L. E. y Casella, G. (1999) dicen hay tres métodos generales para tratar con el sesgo:

1. Modelos de muestreo truncados: asumen que no se publican estudios no significativos.
2. Muestreo invariante: el cual limita el metaanálisis a un subconjunto de estudios que vienen de un marco muestral independiente del proceso de publicación.
3. Aumento de la fuente: se especula sobre el número de estudios perdidos (no vistos) y se ajustan las estimaciones de los efectos en concordancia.

En R existen varios paquetes que permiten realizar metaanálisis, por ejemplo, metafor, mmeta o, desde el punto de vista bayesiano, el paquete bayesmeta (Rover, 2020).

Capítulo 8

Combinación de a priori y análisis preposterior

En el proceso de construcción de la distribución a priori es común enfrentarse al reto de combinar la información previa proveniente de diferentes fuentes, de manera que se obtenga una única distribución que represente la creencia previa sobre el fenómeno de interés. Para lograr esto, se pueden aplicar diferentes técnicas, con las que se espera seleccionar la estrategia más adecuada. Posteriormente, para evaluar la robustez de las distribuciones a priori obtenidas, es necesario realizar un análisis preposterior que permita evaluar su calidad, antes de proceder a la inferencia bayesiana a posteriori.

8.1. Combinación de a priori

En muchas situaciones se presenta tanto la información bibliográfica previa como la presencia de expertos. Debemos desarrollar procedimientos para combinar las a priori provenientes de las dos fuentes. Aparte de esto debe realizarse un análisis para determinar la *compatibilidad* entre ambas fuentes, esto es, determinar hasta qué punto son diferentes. En la mayoría de casos los expertos son conocedores del material bibliográfico disponible, pero poseen información adicional obtenida de otras fuentes que haga más confiable sus creencias.

Si el proceso de elicitación se hace a un grupo de expertos, el investigador cuenta con un conjunto de distribuciones a priori, las cuales deben ser usadas para responder a las hipótesis de interés. En esta situación, diferentes procedimientos están disponibles en la literatura para reducir las distribuciones a priori a una. Winkler (1967b, 1968, 1969) estudió el problema de consenso en el cual las distribuciones a priori son evaluadas por expertos y combinadas en una única distribución para usar esta en un análisis bayesiano posterior. Albert et al. (2007) proponen la combinación de juicios de expertos usando un modelo jerárquico que considera el sesgo y la pre-

cisión individual de los expertos a partir de sus diferencias naturales; luego, Albert et al. (2012) combinan expertos mediante el uso de modelos estadísticos tradicionales.

Note que nadie garantiza que si se tienen diferentes expertos, ellos poseen el mismo grado de experticia y, por lo tanto, sus distribuciones a priori pueden ser mezcladas sin considerar algún tipo de ponderación en sus creencias. Así, al intentar reducir a una única a priori un conjunto de distribuciones elicítadas, se requiere aplicar metodologías, al igual que Albert et al (2007), que combinen dichas creencias de manera adecuada.

Si pensamos en el principio de *la sabiduría de las masas*, y en cómo la elicitación de juicios expertos se relaciona directamente con este principio, haríamos un mayor uso de los métodos de elicitación para ayudar a la toma de decisiones. Entonces, a pesar de la importancia de la opinión de un experto, es preferible elicitar el juicio de múltiples expertos (Jenkinson, 2005).

Existen dos métodos para reducir el número de distribuciones a priori a una única distribución. El primero incluye herramientas matemáticas para obtener la a priori, y el otro permite obtener una única a priori mediante consenso; es ahora cuando el método Delphi, el cual ya comentamos, puede ser aplicado a pesar de que cuando el número de expertos es muy grande, este se vuelve agotador. La formulación matemática más común se hace usando un promedio o un promedio ponderado de las distribuciones elicítadas. Para el caso del promedio ponderado, es necesario caracterizar a los expertos con el fin de hacer una correcta ponderación. Independientemente de lo cuidadosos que seamos en el proceso de elicitación, es recomendable chequear que haya consistencia entre las distribuciones obtenidas y sus creencias iniciales.

Barrera-Causil, et al. (2021) proponen la aplicación de la teoría del análisis de datos funcional para el estudio de las distribuciones a priori provenientes de expertos. Esto abre la posibilidad de adoptar múltiples herramientas estadísticas para el análisis de distribuciones a priori procedentes de juicios de expertos.

De acuerdo con Plous (1993), Hastorf y Cantril (1954) y Loy y Andrews (1981), nuestras creencias pueden ser diferentes aun cuando tengamos el mismo grado de experticia y hayamos experimentado los mismos eventos. Por ejemplo, en un juego de fútbol, las percepciones de dos técnicos igualmente preparados pueden diferir de forma considerable, lo que conlleva desacuerdos para la toma de decisiones. Es por esto que la propuesta de Barrera-Causil, et al. (2021) respecto a la construcción de clústeres de a priori provenientes de expertos, haciendo uso de la teoría del análisis de datos funcional, es una alternativa que ofrece la posibilidad de analizar y hacer inferencia bayesiana de forma separada por grupos de expertos para luego poder concluir sobre la hipótesis de interés. Esta propuesta se describe en la siguiente figura:

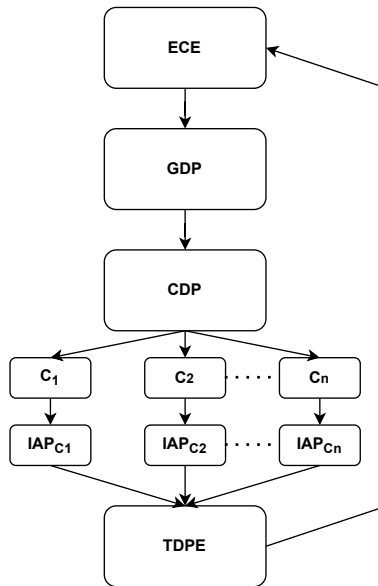


Figura 8.1. Ilustración del flujo para la elicitación de un grupo de expertos

En la figura 8.1 observamos una ilustración del flujo que proponen Barrera-Causil, et al. (2021) para la elicitación de un grupo de expertos. Aquí, ECE =elicitación del conocimiento de expertos, GDP =generación de distribuciones a priori, CDP =clúster de distribuciones a priori, C_i =clúster i de distribuciones a priori, IAP =inferencia a priori o a posteriori, y $TDPE$ =toma de decisiones con panel de expertos. Note que este flujo comprende inicialmente la elicitación de un grupo de expertos, luego se hace la construcción de las distribuciones a priori correspondientes; posteriormente, se construyen clústeres de a priori de expertos y se hace luego inferencia a priori o a posteriori con el ánimo de tomar decisiones con el panel de expertos.

8.2. Análisis preposterior

Dadas las características del proceso bayesiano, es sano tener una posición crítica con relación a cualquier a priori que obtengamos por cualquier método; deberíamos realizar pruebas que permitan determinar, de alguna manera, la calidad de la distribución a priori. Martz y Waller (1982) recomiendan lo siguiente para garantizar un buen análisis bayesiano:

1. Una justificación y análisis detallados de la distribución a priori seleccionada, con un claro entendimiento de las implicaciones matemáticas de esta. Por ejemplo, la selección de una familia normal para representar nuestro conocimiento a priori sobre un parámetro nos restringe a una clase de distribuciones unimodales.
2. Una documentación completa de las fuentes de datos utilizados en la identificación y selección de la a priori.

3. Un análisis preposterior de la distribución a priori con resultados de prueba hipotéticos.
4. Una distribución a posteriori claramente definida para los parámetros de interés.
5. Un análisis de sensibilidad de las inferencias bayesianas para el modelo a priori seleccionado.

Cada uno de estos puntos debe desarrollarse cuidadosamente tanto en el proceso de elicitación como en el de validación de la información recogida. El segundo paso de la lista de Martz y Waller es ejecutado casi simultáneamente en la misma elicitación. Al experto se le alienta a dar respuestas, pero el elicitado debe proporcionar una retroalimentación oportuna, lo que en términos prácticos significa que es inmediata. En la actualidad, con el uso de computadores portátiles, realizar esta tarea no es difícil. Se pueden generar muestras la población con diferentes condiciones planteadas en la a priori; por ejemplo, podemos tomar la mediana, la media, la moda y cuartiles para generar estas muestras. Si el sujeto se siente cómodo con los valores que las muestras están presentado y él cree que no aparecen valores muy raros, o no observa valores que considera deberían estar presentes en la población, esto obliga a revisar la elicitación.

Es necesario documentar detalladamente el proceso, bien sea con grabaciones o registro filmico, para su posterior análisis. Cuando la distribución a priori es obtenida a partir de estudios o datos previos, es necesario justificar plenamente la pertinencia de esos estudios o datos en el problema que estamos enfrentando. Un estudio realizado sobre consumo de drogas en Medellín puede no ser útil para uno que se realice en alguna ciudad del exterior, pero ser pertinente para un estudio similar en Cali. En otros casos puede darse la pertinencia debido a que se estudian fenómenos biológicos que puedan, de alguna manera, ser generalizados a poblaciones de muchos lugares, por ejemplo, si queremos determinar la duración promedio de un embarazo.

La determinación de la distribución a priori puede realizarse de muchas formas, unas más fáciles que otras. De cualquier manera, es recomendable usar varios procedimientos para su elicitación, los cuales deben coincidir de algún modo. La distribución elicitada a mano alzada debería ser muy parecida a algún modelo paramétrico ajustado, por ejemplo.

El análisis preposterior debe realizarse con la distribución a priori para verificar la confianza de los resultados obtenidos como resultantes del proceso elicitado. Por ejemplo, si se elicita el número promedio de goles realizados en un partido de fútbol y, además, se asume que el número de goles marcados en un partido se puede modelar mediante la distribución de Poisson, valores de la probabilidad de observar ciertos marcadores, condicionados en valores tomados de la a priori sobre la media, deben ser consistentes con lo que el sujeto piensa. Este análisis es equivalente a uno

exploratorio de datos (EDA), en el cual se buscan características y problemas con la distribución a priori especificada. Luego de este análisis uno podría responder afirmativamente a la siguiente pregunta: ¿estaría confiado, bajo el supuesto de que no pueda obtener información muestral, para usar esta distribución para realizar todo el trabajo inferencial? En Grabsky (2014) se hace una elicitación en la que se elicitó el tiempo medio de falla. La pregunta, al proponerse directamente sobre la media, puede generar confusión entre los expertos, ya que algunos tienden a pensar más en valores representativos como la moda o como la mediana y, en el caso de distribuciones de sobrevida, estas son en general asimétricas y la media es diferente a la moda y a la mediana, lo cual no ocurre en el caso de distribuciones simétricas.

Otro punto que se analiza en esta etapa es la calidad de la a priori. Para esto, se puede preguntar lo siguiente: ¿es creíble?, ¿se puede utilizar fácilmente a nivel computacional?, ¿está completamente definida sobre el espacio parametral?, ¿cambia radicalmente se se obtienen unos pocos datos?

Uno de los asuntos más delicados es la obtención de dos o más a priori que sean muy diferentes por parte de los expertos o por distintos métodos de elicitación. En caso de diferencias irreconciliables, es sano trabajar paralelamente con varias a priori. Hoy en día esto no es un gran problema, debido a la disponibilidad de recursos computacionales.

8.2.1. Distribución predictiva a priori

La construcción de la distribución predictiva a priori es una de las herramientas que tiene el analista para determinar la calidad de la distribución a priori elicitada, además de entregar información que puede ser útil para determinar los procedimientos de análisis posteriores. Si $\xi(\boldsymbol{\theta})$ representa la distribución a priori sobre $\boldsymbol{\theta}$ elicitada y si se conoce (obviamente hasta cierto nivel) la distribución que genera datos, entonces la distribución predictiva a priori se define como:

$$p(y) = \int f(y|\boldsymbol{\theta}) \xi(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (8.1)$$

Si los datos generados por esta distribución se *acomodan* a lo que el analista cree que deben ser, entonces puede, de cierta manera, pensar que la a priori concuerda con lo que piensa y proceder al proceso posterior.

Este análisis se puede realizar con diferentes distribuciones a priori, de tal forma que se puede mirar la robustez de los resultados ante cambios de la a priori. Observemos que este análisis no es posible de realizarse con distribuciones a priori no informativas.

Número promedios de goles en el fútbol colombiano. Supongamos que elicitamos el número promedio de goles que se marcan en un partido del fútbol profesional en Colombia a un experto y obtenemos que $\theta \sim N(2.5, (0.20)^2)$. Además, si

pensamos que el número de goles que se marcan en un partido del rentado colombiano es $Poisson(\theta)$, la distribución predictiva será:

$$p(y) = \int_0^{\infty} \frac{\theta^y \exp(-\theta)}{y!} \frac{1}{\sqrt{\pi}0.20} \exp\left(-\frac{1}{2 \times 0.20^2}(\theta - 2.5)^2\right) d\theta \quad (8.2)$$

```
# Simulación de la distribución predictiva

# Número promedio de goles del campeonato colombiano
# A priori: normal(2.5, 0.20^2)
# Distribución muestral: Poisson(theta)

dist.predictiva <- function(Nsim = 10000, media = 2.5, dt = 0.20) {
  res <- rpois(Nsim, rnorm(Nsim, mean = media, sd = dt))
  tabla <- table(res)
  print(tabla / sum(tabla))
  barplot(tabla)
  print(summary(res))
  print(var(res))
}

dist.predictiva()
title(main = 'Distribución predictiva\n para el campeonato
colombiano', xlab = 'Número de goles',ylab = 'Frecuencia')
legend(5, 2000,
       legend = c('Media a priori = 2.5', 'Desviación estándar a
priori = 0.20'),
       cex = 0.8, bty = "n")

res
##      0      1      2      3      4      5      6      7      8
##      9
## 0.0829 0.2030 0.2615 0.2120 0.1286 0.0680 0.0303 0.0104 0.0024
## 0.0009
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000  1.000   2.000   2.498   3.000   9.000
## [1] 2.500247
```

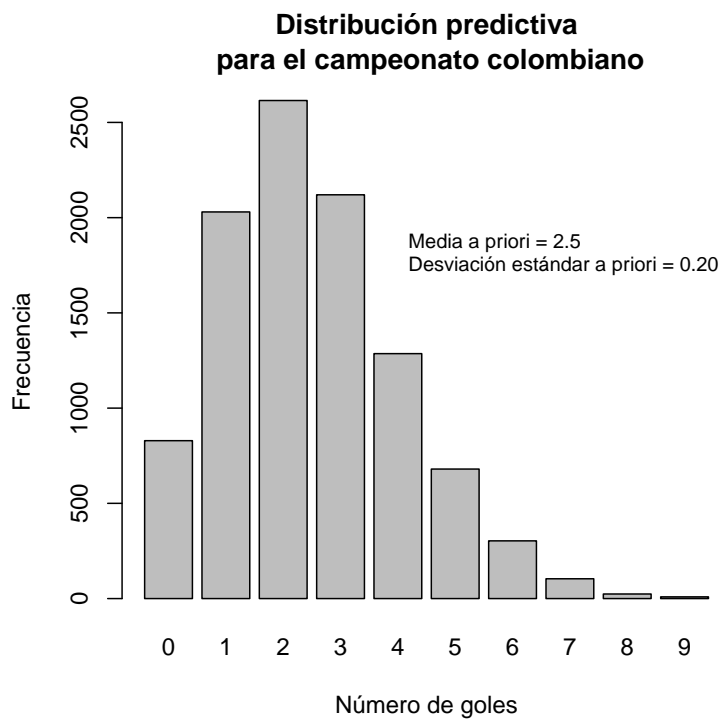


Figura 8.2. *Distribución predictiva a priori*

Capítulo 9

Apéndices

En el siguiente enlace podemos encontrar el repositorio GitHub donde se encuentran los códigos de R usados a lo largo del texto.

https://github.com/cbarrera2101/Elicitacion_Libro.git

9.1. Programas creados para elicitación de distribuciones de probabilidad

9.1.1. Elicitor

James et al. (2010) presentan un programa para elicitar la distribución a priori en regresión bayesiana.

9.1.2. MATCH

Morris et al. (2014) presentan una herramienta en la web para permitir realizar algunas elicitaciones.

9.2. Programas en *R*

```
#####
# Elicitación de la distribución a priori de una proporción
# Se asume que la a priori es una beta(alfa,beta)
# El procedimiento consiste en determinar los límites más extremos
  plausibles, o sea, donde estamos casi convencidos cae la
  verdadera proporción y el valor de la moda.

conseguir.beta.apriori<-
function(){
  # Entrada de la información del elicitor
  LI<-as.numeric(readline("Cuál es el menor valor posible para la
    proporción? "))
  LS<-as.numeric(readline("Cuál es el mayor valor posible para la
    proporción? "))
  moda<-as.numeric(readline("Cuál es el valor más posible para la
    proporción? "))

  funcion.a.mini<-function(parametros,moda=0.5,mini=0.3,maxi=0.6){
    a1<-parametros[1]
    b1<-parametros[2]
    temp<-abs((a1-1)/(a1+b1-2)-moda)+abs(pbeta(mini,a1,b1)-0.01)+abs(
      pbeta(maxi,a1,b1)-0.99)
    return(temp)
  }
  res<-optim(c(2,2),funcion.a.mini,moda=moda,mini=LI,maxi=LS,lower=c
    (1,1),method='L')
  parametros<-res$par
  alfa<-parametros[1]
  beta<-parametros[2]
  x<-seq(0.01,0.99,length=200)
  y<-dbeta(x,alfa,beta)
  plot(x,y,type='l',ylab='',xlab='')
  abline(v=c(LI,moda,LS),lty=2)
  return(res)
}

# Esta función utiliza la aproximación de Perry y Greig (1975)
# Fin de función principal
```

```

conseguir.beta.apriori<-
function(){
  # Entrada de la información del elicitor
  LI<-as.numeric(readline("Cuál es el menor valor posible para la
proporción? "))
  LS<-as.numeric(readline("Cuál es el mayor valor posible para la
proporción? "))
  moda<-as.numeric(readline("Cuál es el valor más posible para la
proporción? "))
  sigma<-(LS-LI)/3.25
  mu<-(LI+0.95*moda+LS)/2.95
  print(mu)
  print(sigma)
  alfa<-(mu^2*(1-mu))/sigma^2-mu
  print(alfa)
  beta<-alfa*(1-mu)/mu
  print(beta)
  x<-seq(0.01,0.99,length=200)
  y<-dbeta(x,alfa,beta)
  plot(x,y,type='l',ylab='',xlab='')
  abline(v=c(LI,moda,LS),lty=2)
  abline(h=0)
  list(alfa=alfa,beta=beta,media=mu,desvi.tipica=sigma)
}

# Fin de función principal

resultados<-conseguir.beta.apriori()

resultados
alfa0<-resultados$par[1]
beta0<-resultados$par[2]
x<-seq(0,1,length=1000)
y<-dbeta(x,alfa0,beta0)
par(mfrow=c(3,1))
plot(x,y,ylab='Densidad',xlab=expression(pi),type='l')
title(main='Distribución a priori')
# Saco una muestra de 10 estudiantes al azar y obtengo 2 mujeres.
y2<-x^2*(1-x)^8
# Verosimilitud
plot(x,y2,type='l',col='red',ylab='',xlab='')
title(main='Función de Verosimilitud')
# a posteriori
y3<-dbeta(x,alfa0+2,beta0+8)
plot(x,y3,ylab='Densidad',xlab=expression(pi),type='l')
title(main='Distribución a posteriori')

```

Bibliografía

- Adams, F. K. (2006). Expert elicitation and Bayesian analysis of construction contract risks: an investigation. *Construction Management and Economics*, *24*(1), 81-96. <https://doi.org/10.1080/01446190500310254>
- Al-Awadhi, S. & Garthwaite, P. (2001). Prior distribution assessment for a multivariate normal distribution: an experimental study. *Journal of Applied Statistics*, *28*(1), 5-23. <https://doi.org/10.1080/02664760120011788>
- Al-Awadhi, S. A. & Garthwaite, P. H. (1998). An elicitation method for multivariate normal distributions. *Communications in Statistics-Theory and Methods*, *27*(5), 1123-1142. <https://doi.org/10.1080/03610929808832149>
- Al-Awadhi, S. A. & Garthwaite, P. H. (2006). Quantifying expert opinion for modelling fauna habitat distributions. *Computational Statistics*, *21*(1), 121-140. <https://doi.org/10.1007/s00180-006-0255-x>
- Albert, I., Donnet, S., Guihenneuc-Jouyaux, C., Low-Choy, S., Mengersen, K. & Rousseau, J. (2012a). Combining expert opinions in prior elicitation. *Bayesian Analysis*, *7*(3), 503-532. <https://doi.org/10.1214/12-BA717>
- Albert, I., Donnet, S., Guihenneuc-Jouyaux, C., Low-Choy, S., Mengersen, K. & Rousseau, J. (2012b). Combining expert opinions in prior elicitation. *Bayesian Analysis*, *7*(3), 503-532. <https://doi.org/10.1214/12-BA717>
- Anscombe, F. J., Aumann, R. J. y col. (1963). A definition of subjective probability. *Annals of mathematical statistics*, *34*(1), 199-205. <https://doi.org/10.1214/aoms/1177704255>
- Arkes, H. R., Mumpower, J. L. & Stewart, T. R. (1997). Combining expert opinions. *Science*, *275*(5299), 461-465. <https://doi.org/10.1126/science.275.5299.461>
- Ayyub, B. M. (2000). Methods for expert-opinion elicitation of probabilities and consequences for Corps facilities. *US Army Corps of Engineers, IWR Report No. 00-R-10*. <https://www.iwr.usace.army.mil/portals/70/docs/iwrreports/00-r-101.pdf>
- Ayyub, B. M. (2001a). *Elicitation of expert opinions for uncertainty and risks*. CRC press. <https://doi.org/10.1201/9781420040906>
- Ayyub, B. M. (2001b). A practical guide on conducting expert-opinion elicitation of probabilities and consequences for corps facilities. *Institute for Water Resources, Alexandria, VA, USA*. <https://www.iwr.usace.army.mil/Portals/70/docs/iwrreports/01-R-01.pdf>

- Barrera-Causil, C., Correa, J. C., Zamecnik, A., Torres-Avilés, F. & Marmolejo-Ramos, F. (2021). An FDA-Based Approach for Clustering Elicited Expert Knowledge. *Stats*, 4(1), 184-204. <https://doi.org/10.3390/stats4010014>
- Barrera-Causil, C. J., Correa, J. C. & Marmolejo-Ramos, F. (2019). Experimental investigation on the elicitation of subjective distributions. *Frontiers in Psychology*, 10, 862. <https://doi.org/10.3389/fpsyg.2019.00862>
- Bassett, W. F. & Lumsdaine, R. L. (2001). Probability limits: Are subjective assessments adequately accurate? *Journal of Human Resources*, 327-363. <https://doi.org/10.2307/3069662>
- Benson, P. G., Curley, S. P. & Smith, G. F. (1995). Belief assessment: An underdeveloped phase of probability elicitation. *Management Science*, 41(10), 1639-1653. <https://doi.org/10.1287/mnsc.41.10.1639>
- Biedermann, A., Bozza, S., Taroni, F. & Aitken, C. (2017). The consequences of understanding expert probability reporting as a decision. *Science & Justice*, 57(1), 80-85. <https://doi.org/10.1016/j.scijus.2016.08.001>
- Biedermann, A., Garbolino, P. & Taroni, F. (2013). The subjectivist interpretation of probability and the problem of individualisation in forensic science. *Science & Justice*, 53(2), 192-200. <https://doi.org/10.1016/j.scijus.2012.09.003>
- Bleichrodt, H. & Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management science*, 46(11), 1485-1496. <https://doi.org/10.1287/mnsc.46.11.1485.12035>
- Bolado, R., Duro, L., Cormenzana, J., Cuñado, M., Grivé, M. & Bennett, D. (2009). *Expert judgement review and exercise (estimation of the solubility limit distributions for five elements in the near field in a repository in granite)* MILESTONE (MN: 2.2. A. 12) (inf. téc.). MILESTONE (MN: 2.2. A. 12). <http://www.ip-pamina.eu/downloads/pamina.m2.2.a.12.pdf>
- Bormane, L., Gržibovska, J., Bērziša, S. & Grabis, J. (2016). Impact of requirements elicitation processes on success of information system development projects. *Information Technology and Management Science*, 19(1), 57-64. <https://doi.org/10.1515/itms-2016-0008>
- Bottomley, P. A. & Doyle, J. R. (2001). A comparison of three weight elicitation methods: good, better, and best. *Omega*, 29(6), 553-560. [https://doi.org/10.1016/S0305-0483\(01\)00043-7](https://doi.org/10.1016/S0305-0483(01)00043-7)
- Bottomley, P. A., Doyle, J. R. & Green, R. H. (2000). Testing the reliability of weight elicitation methods: direct rating versus point allocation. *Journal of Marketing Research*, 37(4), 508-513. <https://doi.org/10.1509/jmkr.37.4.508.18737>
- Bravo, L. E., Cortés, A., Carrascal, E., Jaramillo, R., Garcíea, L. S., Bravo, P. E., Badel, A. & Bravo, P. A. (2003). Helicobacter pylori: patología y prevalencia en biopsias gástricas en Colombia. *Colombia médica*, 34(3), 124-131. <https://doi.org/10.25100/cm.v34i3.263>
- Brenner, L. A., Koehler, D. J., Liberman, V. & Tversky, A. (1996). Overconfidence in probability and frequency judgments: A critical examination. *Organizational*

- Behavior and Human Decision Processes*, 65(3), 212-219. <https://doi.org/10.1006/obhd.1996.0021>
- Brockwell, S. E. & Gordon, I. R. (2001). A comparison of statistical methods for meta-analysis. *Statistics in medicine*, 20(6), 825-840. <https://doi.org/10.1002/sim.650>
- Brookshire, D. S. & Coursey, D. L. (1987). Measuring the value of a public good: an empirical comparison of elicitation procedures. *The American Economic Review*, 554-566. <https://www.jstor.org/stable/1814530>
- Brown, B. B. (1968). *Delphi process: a methodology used for the elicitation of opinions of experts* (inf. téc.). Rand Corp Santa Monica CA. <https://doi.org/10.7249/P3925>
- Chacko, G. K. (1991). *Decision-making under uncertainty: an applied statistics approach*. Praeger Pub Text. <https://researchdiscovery.drexel.edu/esploro/outputs/bookReview/Decision-making-under-uncertainty---an-applied/991019238689804721>
- Chaloner, K., Church, T., Louis, T. A. & Matts, J. P. (1993). Graphical elicitation of a prior distribution for a clinical trial. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 42(4), 341-353. <https://doi.org/10.2307/2348670>
- Chaloner, K. M. & Duncan, G. T. (1983). Assessment of a beta prior distribution: PM elicitation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 32(1-2), 174-180. <https://doi.org/10.2307/2987588>
- Chen, M.-H., Ibrahim, J. G., Shao, Q.-M. & Weiss, R. E. (2003). Prior elicitation for model selection and estimation in generalized linear mixed models. *Journal of Statistical Planning and Inference*, 111(1-2), 57-76. [https://doi.org/10.1016/S0378-3758\(02\)00294-5](https://doi.org/10.1016/S0378-3758(02)00294-5)
- Chesley, G. (1975). Elicitation of subjective probabilities: A review. *The Accounting Review*, 50(2), 325-337. <https://doi.org/10.2307/245498>
- Chesley, G. (1976). The elicitation of subjective probabilities: A laboratory study in an accounting context. *Journal of Accounting Research*, 27-48. <https://doi.org/10.2307/2490420>
- Chesley, G. (1977). Subjective probability elicitation: The effect of congruity of datum and response mode on performance. *Journal of Accounting Research*, 1-11. <https://doi.org/10.2307/2490484>
- Chesley, G. (1978). Subjective probability elicitation techniques: A performance comparison. *Journal of Accounting Research*, 225-241. <https://doi.org/10.2307/2490655>
- Choy, S. L., O'Leary, R. & Mengersen, K. (2009). Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. *Ecology*, 90(1), 265-277. <https://doi.org/10.1890/07-1886.1>
- Christov, S. C., Marquard, J. L., Avrunin, G. S. & Clarke, L. A. (2017). Assessing the effectiveness of five process elicitation methods: A case study of chemotherapy treatment plan review. *Applied Ergonomics*, 59, 364-376. <https://doi.org/10.1016/j.apergo.2016.10.017>

- Clemen, R. T. & Winkler, R. L. (1990). Unanimity and compromise among probability forecasters. *Management science*, 36(7), 767-779. <https://doi.org/10.1287/mnsc.36.7.767>
- Clemen, R. T. & Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk analysis*, 19(2), 187-203. <https://doi.org/10.1023/A:1006917509560>
- Cooke, R. y col. (1991). *Experts in uncertainty: opinion and subjective probability in science*. Oxford University Press on Demand. <https://doi.org/10.1093/oso/9780195064650.001.0001>
- Cooley, C. A. & MacEachern, S. N. (1999). Prior elicitation in the classification problem. *Canadian Journal of Statistics*, 27(2), 299-313. <https://doi.org/10.2307/3315980>
- Corless, J. C. (1972). Assessing prior distributions for applying Bayesian statistics in auditing. *The Accounting Review*, 47(3), 556-566. <https://doi.org/10.2307/245206>
- Correa-Morales, J. C. & Barrera-Causil, C. (2021). Elicitation of the Parameters of Multiple Linear Models. *Revista Colombiana de Estadística*, 44(1), 159-170. <https://doi.org/10.15446/rce.v44n1.90549>
- Cressey, P., Lake, R. & Thornley, C. (2016). *Expert elicitation-derived estimates of foodborne proportions of enteric illnesses* (inf. téc.). Ministry for Primary Industries. MPI Technical Report – Paper No: 2016/38, New Zeland. <https://www.mpi.govt.nz/dmsdocument/12936-Expert-elicitation-derived-estimates-of-foodborne-proportions-of-enteric-illnessesMPI>
- Crosby, M. A. (1980). Implications of prior probability elicitation on auditor sample size decisions. *Journal of Accounting Research*, 585-593. <https://doi.org/10.2307/2490477>
- Crosby, M. A. (1981). Bayesian statistics in auditing: a comparison of probability elicitation techniques. *Accounting Review*, 355-365. <https://doi.org/10.2307/246679>
- DeGroot, M. H. (2005). *Optimal statistical decisions*. John Wiley & Sons. <https://doi.org/10.1002/0471722065>
- Del Re, A. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, 11(1), 37-50. <https://doi.org/10.20982/tqmp.11.1.p037>
- Delavande, A., Giné, X. & McKenzie, D. (2011). Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design? *Journal of Applied Econometrics*, 26(3), 479-497. <https://doi.org/10.1002/jae.1222>
- Demuyne, T. (2013). A mechanism for eliciting the mean and quantiles of a random variable. *Economics letters*, 121(1), 121-123. <https://doi.org/10.1016/j.econlet.2013.07.019>
- Devilee, J. & Knol, A. (2012). *Software to support expert elicitation: An exploratory study of existing software packages* (inf. téc.). Rijksinstituut voor Volksge-

- zondheid en Milieu RIVM. <https://rivm.openrepository.com/handle/10029/259480>
- Dewispelare, A. R., Herren, L. T. & Clemen, R. T. (1995). The use of probability elicitation in the high-level nuclear waste regulation program. *International Journal of Forecasting*, 11(1), 5-24. [https://doi.org/10.1016/0169-2070\(94\)02006-B](https://doi.org/10.1016/0169-2070(94)02006-B)
- Dominici, F., Parmigiani, G. & Clyde, M. (2000). Conjugate analysis of multivariate normal data with incomplete observations. *Canadian Journal of Statistics*, 28(3), 533-550. <https://doi.org/10.2307/3315963>
- Dominitz, J. & Manski, C. (1996). Eliciting Student Expectations of the Returns to Schooling. *Journal of Human Resources*, 31(1), 1-26. <https://doi.org/10.2307/146041>
- Druzdzel, M. J. (1989). Verbal uncertainty expressions: Literature review. *Pittsburgh, PA: Carnegie Mellon University, Department of Engineering and Public Policy*, 1-13. <https://sites.pitt.edu/~druzdzel/abstracts/verbal.html>
- Durbach, I., Merven, B. & McCall, B. (2017). Expert elicitation of autocorrelated time series with application to e3 (energy-environment-economic) forecasting models. *Environmental Modelling & Software*, 88, 93-105. <https://doi.org/10.1016/j.envsoft.2016.11.007>
- Eberly, L. E. & Casella, G. (1999). Bayesian estimation of the number of unseen studies in a meta-analysis. *Journal of official statistics-stockholm*, 15(4), 477-494. <https://doi.org/10.1002/sim.9695>
- Fauss, E. K. (2008). *Risk Analysis of Nanotechnology through Expert Elicitation: A Silver Nanotechnology Case Study* (Tesis doctoral). University of Virginia. https://wecanfigurethisout.org/NANO/lecture_notes/Nano_challenges_and_fears_Supporting_materials_files/Nanoparticle_Ag_toxicity/Risk%20analysis%20of%20silver%20nanotechnology%20-%20MS%20thesis%20-%20Emma%20Fauss.pdf
- Ferson, S., O'Rawe, J., Antonenko, A., Siegrist, J., Mickley, J., Luhmann, C. C., Sentz, K. & Finkel, A. M. (2015). Natural language of uncertainty: numeric hedge words. *International Journal of Approximate Reasoning*, 57, 19-39. <https://doi.org/10.1016/j.ijar.2014.11.003>
- Fillenbaum, S., Wallsten, T. S., Cohen, B. L. & Cox, J. A. (1991). Some effects of vocabulary and communication task on the understanding and use of vague probability expressions. *The American Journal of Psychology*, 35-60. <https://doi.org/10.2307/1422850>
- Firmino, P. R. A. & Droguett, E. L. (2015). An expert opinion elicitation method based on binary search and Bayesian intervals. *optimization*, 45, 46. <https://doi.org/10.1504/IJRAM.2015.071224>
- Flórez, A. F., Correa, J. C., Garcíea, M. y col. (2015). Interactive R App for Carrying Out an Elicitation Process Applied to Prostate Cancer in Colombia. *Romanian Statistical Review*, 63(2), 119-129. https://www.revistadestatistica.ro/wp-content/uploads/2015/04/RRS2_2015_A12.pdf

- Flórez, A. F. & Morales, J. C. C. (2015). Una propuesta metodológica para elicitar el vector de parámetros π de la distribución Multinomial. *Comunicaciones en Estadística*, 8(1), 81-97. <https://doi.org/10.15332/s2027-3355.2015.0001.05>
- Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczkowski, A. & Wreathall, J. (2004). Expert elicitation approach for performing ATHEANA quantification. *Reliability Engineering & System Safety*, 83(2), 207-220. <https://doi.org/10.1016/j.ress.2003.09.011>
- Franklin, J. (2001). *The science of conjecture: Evidence and probability before Pascal*. Johns Hopkins University Press. https://books.google.com.co/books?id=_9k-4DJhK-wC
- Galway, L. A. (2007). *Subjective probability distribution elicitation in cost risk analysis: A review* (inf. téc.). Rand Corporation. <https://doi.org/10.7249/TR410>
- Garthwaite, P. H. (1994). Assessment of prior distributions for regression models: an experimental study. *Communications in Statistics-Simulation and Computation*, 23(4), 871-895. <https://doi.org/10.1080/03610919408813206>
- Garthwaite, P. H. & Al-Awadhi, S. A. (2001). Non-conjugate prior distribution assessment for multivariate normal sampling. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(1), 95-110. <https://doi.org/10.1111/1467-9868.00278>
- Garthwaite, P. H. & Dickey, J. M. (1988). Quantifying expert opinion in linear regression problems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(3), 462-474. <https://doi.org/10.1111/j.2517-6161.1988.tb01741.x>
- Garthwaite, P. H. & Dickey, J. M. (1992). Elicitation of prior distributions for variable-selection problems in regression. *The Annals of Statistics*, 20(4), 1697-1719. <https://doi.org/10.1214/aos/1176348886>
- Garthwaite, P. H., Jenkinson, D. J., Rakow, T. & Wang, D. D. (2007). *Comparison of fixed and variable interval methods for eliciting subjective probability distributions* (inf. téc.). Tech. rep., University of New South Wales. <https://www.stat.cmu.edu/tr/tr808/tr808.pdf>
- Garthwaite, P. H., Kadane, J. B. & O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American statistical Association*, 100(470), 680-701. <https://doi.org/10.1198/016214505000000105>
- Garthwaite, P., Kadane, J. & O'Hagan, A. (2005). Elicitation. <https://tinyurl.com/bdhpz7m5>
- Gavasakar, U. (1988). A comparison of two elicitation methods for a prior distribution for a binomial parameter. *Management Science*, 34(6), 784-790. <https://doi.org/10.1287/mnsc.34.6.784>
- Gelfand, A. E., Mallick, B. K. & Dey, D. K. (1995). Modeling expert opinion arising as a partial probabilistic specification. *Journal of the American Statistical Association*, 90(430), 598-604. <https://doi.org/10.1080/01621459.1995.10476552>

- Genest, C. & Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1), 114-135. <https://doi.org/10.1214/ss/1177013830>
- Gigerenzer, G. & Garcia-Retamero, R. (2017). Cassandra's regret: The psychology of not wanting to know. *Psychological review*, 124(2), 179. <https://doi.org/10.1037/rev0000055>
- Gigerenzer, G., Hoffrage, U. & Kleinbölting, H. (1991). Probabilistic mental models: a Brunswikian theory of confidence. *Psychological review*, 98(4), 506. <https://doi.org/10.1037/0033-295x.98.4.506>
- Gill, J. & Walker, L. D. (2005). Elicited priors for Bayesian model specifications in political science research. *The Journal of Politics*, 67(3), 841-872. <https://doi.org/10.1111/j.1468-2508.2005.00342.x>
- Gonzalez-Vallejo, C. C., Erev, I. & Wallsten, T. S. (1994). Do decision quality and preference order depend on whether probabilities are verbal or numerical? *The American journal of psychology*, 157-172. <https://doi.org/10.2307/1423035>
- Goulet, V., Jacques, M. & Pigeon, M. (2009). Expert: Modeling Without Data Using Expert Opinion. *R J.*, 1(1), 31. <https://doi.org/10.32614/RJ-2009-005>
- Grabski, F. (2014). Subjective priors in Bayesian estimation of safety and reliability characteristics. *Procedia Engineering*, 84, 62-69. <https://doi.org/10.1016/j.proeng.2014.10.410>
- Greenland, S. (2001). Putting background information about relative risks into conjugate prior distributions. *Biometrics*, 57(3), 663-670. <https://doi.org/10.1111/j.0006-341x.2001.00663.x>
- Grigore, B., Peters, J., Hyde, C. & Stein, K. (2017). EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. *BMC medical informatics and decision making*, 17(1), 1-10. <https://doi.org/10.1186/s12911-017-0527-0>
- Grisley, W. & Kellogg, E. D. (1983). Farmers' subjective probabilities in Northern Thailand: an elicitation analysis. *American Journal of Agricultural Economics*, 65(1), 74-82. <https://doi.org/10.2307/1240339>
- Gupta, A. & Upadhyay, S. (2019). Subjective Elicitation of Dirichlet Hyperparameters Using Past Data: A Study of Ovarian Cancer Patients. *Austrian Journal of Statistics*, 48(3), 1-14. <https://doi.org/10.17713/ajs.v48i3.814>
- Gzyl, H., ter Horst, E. & Molina, G. (2017). Inferring probability densities from expert opinion. *Applied Mathematical Modelling*, 43, 306-320. <https://doi.org/10.1016/j.apm.2016.11.006>
- Hahn, E. D. (2006). Re-examining informative prior elicitation through the lens of Markov chain Monte Carlo methods. *Journal of the royal statistical society: Series A (Statistics in Society)*, 169(1), 37-48. <https://doi.org/10.1111/j.1467-985X.2005.00381.x>
- Hamada, M. S., Martz, H. F., Reese, C. S. & Wilson, A. G. (2008). *Bayesian reliability* (Vol. 15). Springer. <https://link.springer.com/book/10.1007/978-0-387-77950-8>

- Harrison, G. W., Martiénez-Correa, J. & Swarthout, J. T. (2014). Eliciting subjective probabilities with binary lotteries. *Journal of Economic Behavior & Organization*, 101, 128-140. <https://doi.org/10.1016/j.jebo.2014.02.011>
- Hastorf, A. H. & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology*, 49, 129-134. <https://doi.org/10.1037/h0057880>
- Hendrickson, A. D. & Buehler, R. J. (1972). Elicitation of subjective probabilities by sequential choices. *Journal of the American Statistical Association*, 67(340), 880-883. <https://doi.org/10.1080/01621459.1972.10481310>
- Holloway, C. A. (1979). *Decision making under uncertainty: models and choices*. Prentice Hall. <https://books.google.com.co/books?id=ZhFPAAAAMAAJ>
- Hosack, G. R., Hayes, K. R. & Barry, S. C. (2017). Prior elicitation for Bayesian generalised linear models with application to risk control option assessment. *Reliability Engineering & System Safety*, 167, 351-361. <https://doi.org/10.1016/j.ress.2017.06.011>
- Huson, L. & Kinnersley, N. (2009). Bayesian fitting of a logistic dose-response curve with numerically derived priors. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, 8(4), 279-286. <https://doi.org/10.1002/pst.348>
- James, A., Choy, S. L. & Mengersen, K. (2010). Elicitor: an expert elicitation tool for regression in ecology. *Environmental Modelling & Software*, 25(1), 129-145. <https://doi.org/10.1016/j.envsoft.2009.07.003>
- Jenkinson, D. (2005). The elicitation of probabilities - A review of the statistical literature. *BEEP Working Paper, University of Sheffield*. <https://api.semanticscholar.org/CorpusID:16772099>
- Johnson, S. R., Tomlinson, G. A., Hawker, G. A., Granton, J. T., Grosbein, H. A. & Feldman, B. M. (2010). A valid and reliable belief elicitation method for Bayesian priors. *Journal of clinical epidemiology*, 63(4), 370-383. <https://doi.org/10.1016/j.jclinepi.2009.08.005>
- Jolson, M. A. & Rossow, G. L. (1971). The Delphi process in marketing decision making. *Journal of Marketing Research*, 8(4), 443-448. <https://doi.org/10.1177/002224377100800406>
- Jones, G. & Johnson, W. O. (2014). Prior elicitation: Interactive spreadsheet graphics with sliders can be fun, and informative. *The American Statistician*, 68(1), 42-51. <https://doi.org/10.1080/00031305.2013.868828>
- Kadane, J. & Wolfson, L. (1998). Experiences in Elicitation. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47(1), 3-19. <https://doi.org/10.1111/1467-9884.00113>
- Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S. & Peters, S. C. (1980). Interactive elicitation of opinion for a normal linear model. *Journal of the American Statistical Association*, 75(372), 845-854. <https://doi.org/10.1080/01621459.1980.10477562>

- Kadane, J. B. & Winkler, R. L. (1988). Separating probability elicitation from utilities. *Journal of the American Statistical Association*, 83(402), 357-363. <https://doi.org/10.1080/01621459.1988.10478605>
- Karni, E. (1999). Elicitation of subjective probabilities when preferences are state-dependent. *International Economic Review*, 479-486. <https://doi.org/10.1111/1468-2354.00025>
- Kausar, S., Tariq, S., Riaz, S. & Khanum, A. (2010). Guidelines for the selection of elicitation techniques. *2010 6th International Conference on Emerging Technologies (ICET)*, 265-269. <https://doi.org/10.1109/ICET.2010.5638476>
- Keeney, R. L. & Von Winterfeldt, D. (1991). Eliciting probabilities from experts in complex technical problems. *IEEE Transactions on engineering management*, 38(3), 191-201. <https://doi.org/10.1109/17.83752>
- Keeney, S., McKenna, H. & Hasson, F. (2011). *The Delphi technique in nursing and health research*. John Wiley & Sons. <https://doi.org/10.1002/9781444392029>
- Kennedy, P. (1986). The Bayesian approach to research in economic education. *The Journal of Economic Education*, 17(1), 9-24. <https://doi.org/10.1080/00220485.1986.10845142>
- Kinnersley, N. & Day, S. (2013). Structured approach to the elicitation of expert beliefs for a Bayesian-designed clinical trial: a case study. *Pharmaceutical statistics*, 12(2), 104-113. <https://doi.org/10.1002/pst.1552>
- Kloprogge, P., van der Sluijs, J. P., Wardekker, J. A. y col. (2007). *Uncertainty communication: issues and good practice*. Copernicus Institute for Sustainable Development; Innovation. https://www.nusap.net/downloads/reports/uncertainty_communication.pdf
- Knol, A. B., Slottje, P., van der Sluijs, J. P. & Lebet, E. (2010). The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health*, 9(1), 1-16. <https://doi.org/10.1186/1476-069X-9-19>
- Konold, C. (1989). Informal conceptions of probability. *Cognition and instruction*, 6(1), 59-98. https://doi.org/10.1207/s1532690xci0601_3
- Kynn, M. (2008). The “heuristics and biases” bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1), 239-264. <https://doi.org/10.1111/j.1467-985X.2007.00499.x>
- Landeta, J. (1999). *El método Delphi: una técnica de previsión para la incertidumbre*. Ariel Barcelona. <https://dialnet.unirioja.es/servlet/libro?codigo=208626>
- Landquist, H., Norrman, J., Lindhe, A., Norberg, T., Hassellöv, I.-M., Lindgren, J. & Rosen, L. (2017). Expert elicitation for deriving input data for probabilistic risk assessment of shipwrecks. *Marine pollution bulletin*, 125(1-2), 399-415. <https://doi.org/10.1016/j.marpolbul.2017.09.043>
- Lapin, L. L. (1994). *Quantitative methods for business decisions: with cases*. Wadsworth Publishing Company. <https://books.google.com.co/books?id=fl-1AAAAIAAJ>

- Lau, H.-S., Lau, A. H.-L. & Kottas, J. F. (1999). Using Tocher's curve to convert subjective quantile-estimates into a probability distribution function. *IIE transactions*, 31(3), 245-254. <https://doi.org/10.1023/A:1007649822390>
- Leamer, E. E. (1992). Bayesian elicitation diagnostics. *Econometrica: Journal of the Econometric Society*, 919-942. [https://doi.org/0012-9682\(199207\)60:4<919:BED>2.0.CO;2-V](https://doi.org/0012-9682(199207)60:4<919:BED>2.0.CO;2-V)
- Lindley, D. (1983). Reconciliation of probability distributions. *Operations Research*, 31(5), 866-880. <https://doi.org/10.1287/opre.31.5.866>
- Lindley, D. V. (2013). *Understanding uncertainty*. John Wiley & Sons. <https://www.al-edu.com/wp-content/uploads/2014/05/Lindley-D.V.-Understanding-uncertainty-2006.pdf>
- Lindley, D. V. & Singpurwalla, N. D. (1986). Multivariate distributions for the life lengths of components of a system sharing a common environment. *Journal of Applied Probability*, 23(2), 418-431. <https://doi.org/10.2307/3214184>
- List, J. A. (2001). Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards. *American economic review*, 91(5), 1498-1507. <https://doi.org/10.1257/aer.91.5.1498>
- Low Choy, S., James, A. & Mengersen, K. (2009). Expert elicitation and its interface with technology: a review with a view to designing Elicitor. *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand. Cairns, Australia*. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=f4159b2e9cce84345987468d7ed8b5392343683e>
- Loy, J. W. & Andrews, D. S. (1981). They also sew a game: A replication of a case study. *Replications in Social Psychology*, 1, 45-49. <https://cir.nii.ac.jp/crid/1130000794687985152>
- Ludke, R. L., Stauss, F. F. & Gustafson, D. H. (1977). Comparison of five methods for estimating subjective probability distributions. *Organizational Behavior and Human Performance*, 19(1), 162-179. [https://doi.org/10.1016/0030-5073\(77\)90060-5](https://doi.org/10.1016/0030-5073(77)90060-5)
- Luo, S., Chen, Y., Su, X. & Chu, H. (2014). mmeta: an R package for multivariate meta-analysis. *Journal of statistical software*, 56(11), 11. <https://doi.org/10.18637/jss.v056.i11>
- Makridakis, S., Wheelwright, S. C. & Hyndman, R. J. (2008). *Forecasting methods and applications*. John Wiley & Sons. <https://books.google.com.co/books?id=0n4eAQAAIAAJ>
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5), 1329-1376. <https://www.jstor.org/stable/3598892>
- Matheson, J. E. & Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management science*, 22(10), 1087-1096. <https://doi.org/10.1287/mnsc.22.10.1087>
- Mazucchi, T., Cooke, R., Dekker, R., van Dorp, J. & van Noortwijk, J. (1992). *The elicitation and use of expert judgment for maintenance optimization* (inf. téc.).

- The world congress on expert systems proceedings 1991. <http://hdl.handle.net/1765/2207>
- McBride, M. F. & Burgman, M. A. (2012). What is expert knowledge, how is such knowledge gathered, and how do we use it to address questions in landscape ecology? *Expert knowledge and its application in landscape ecology* (pp. 11-38). Springer. https://doi.org/10.1007/978-1-4614-1034-8_2
- Meeden, G. (1992). An elicitation procedure using piecewise conjugate priors. *Bayesian analysis in statistics and econometrics* (pp. 195-206). Springer. https://doi.org/10.1007/978-1-4612-2944-5_11
- Moala, F. A. & O'Hagan, A. (2010). Elicitation of multivariate prior distributions: A nonparametric Bayesian approach. *Journal of Statistical Planning and Inference*, *140*(7), 1635-1655. <https://doi.org/10.1016/j.jspi.2010.01.004>
- Moder, J. J. & Rodgers, E. (1968). Judgment estimates of the moments of PERT type distributions. *Management Science*, *15*(2), B-76. <https://doi.org/10.1287/mnsc.15.2.B76>
- Montoya, Y. & Correa Morales, J. C. (2017). Elicitación del vector de parámetros de la distribución multinomial a partir de varios expertos. *Comunicaciones en Estadística*, *10*(2), 207-223. <https://doi.org/10.15332/2422474x.2884>
- Moore, P. G. (1977). The manager's struggles with uncertainty. *Journal of the Royal Statistical Society: Series A (General)*, *140*(2), 129-148. <https://doi.org/10.2307/2344872>
- Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National academy of Sciences*, *111*(20), 7176-7184. <https://doi.org/10.1073/pnas.1319946111>
- Morris, D. E., Oakley, J. E. & Crowe, J. A. (2014). A web-based tool for eliciting probability distributions from experts. *Environmental Modelling & Software*, *52*, 1-4. <https://doi.org/10.1016/j.envsoft.2013.10.010>
- Morton-Williams, J. (1993). *Interviewer approaches*. Dartmouth. <https://archive.org/details/interviewerappro0000mort>
- Mosteller, F. & Youtz, C. (1990). Quantifying probabilistic expressions. *Statistical Science*, *5*(1), 2-34. <https://www.jstor.org/stable/2245869>
- Nelson, R. G. & Bessler, D. A. (1989). Subjective probabilities and scoring rules: Experimental evidence. *American Journal of Agricultural Economics*, *71*(2), 363-369. <https://doi.org/10.2307/1241594>
- Nemet, G. F., Baker, E. & Jenni, K. E. (2013). Modeling the future costs of carbon capture using experts' elicited probabilities under policy scenarios. *Energy*, *56*, 218-228. <https://doi.org/10.1016/j.energy.2013.04.047>
- Oakley, J. (2010). Eliciting univariate probability distributions. *Rethinking risk measurement and reporting*, *1*, 155-177. <https://tinyurl.com/4knut6ya>
- Oakley, J. E. & O'Hagan, A. (2007). Uncertainty in prior elicitation: a nonparametric approach. *Biometrika*, *94*(2), 427-441. <https://doi.org/10.1093/biomet/asm031>

- O'Connor, M., Remus, W. & Lim, K. (2005). Improving judgmental forecasts with judgmental bootstrapping and task feedback support. *Journal of Behavioral Decision Making*, 18(4), 247-260. <https://doi.org/10.1002/bdm.499>
- O'Hagan, A. (1998). Eliciting expert beliefs in substantial practical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 21-35. <https://doi.org/10.1111/1467-9884.00114>
- O'Hagan, A. & Oakley, J. E. (2004). Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety*, 85(1-3), 239-248. <https://doi.org/10.1016/j.ress.2004.03.014>
- O'Leary, R., Mengersen, K., Murray, J. & Choy, S. L. (2009). Comparison of four expert elicitation methods: For Bayesian logistic regression and classification trees. *18th World IMACS/MODSIM congress*. <https://espace.library.uq.edu.au/view/UQ:316810>
- O'Leary, R. A., Choy, S. L., Murray, J. V., Kynn, M., Denham, R., Martin, T. G. & Mengersen, K. (2009). Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby *Petrogale penicillata*. *Environmetrics: The official journal of the International Environmetrics Society*, 20(4), 379-398. <https://doi.org/10.1002/env.935>
- Papathomas, M. & Hocking, R. J. (2003). Bayesian updating for binary variables: an application in the UK water industry. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(4), 483-499. <https://doi.org/10.1046/j.0039-0526.2003.00451.x>
- Perry, C. & Greig, I. (1975). Estimating the mean and variance of subjective distributions in PERT and decision analysis. *Management Science*, 21(12), 1477-1480. <https://doi.org/10.1287/mnsc.21.12.1477>
- Peters, U., Poole, C. & Arab, L. (2001). Does tea affect cardiovascular disease? A meta-analysis. *American Journal of Epidemiology*, 154(6), 495-503. <https://doi.org/10.1093/aje/154.6.495>
- Petitti, D. B. (2001). Approaches to heterogeneity in meta-analysis. *Statistics in medicine*, 20(23), 3625-3633. <https://doi.org/10.1002/sim.1091>
- Plous, S. (1993). *The Psychology of Judgment and Decision Making*. McGraw-Hill, Inc.
- Poirier, D. J. (1988). Frequentist and subjectivist perspectives on the problems of model building in economics. *Journal of economic perspectives*, 2(1), 121-144. <https://doi.org/10.1257/jep.2.1.121>
- Poirier, D. J. (1998). Revising beliefs in nonidentified models. *Econometric theory*, 14(4), 483-509. <https://doi.org/10.1017/S0266466698144043>
- Pratt, J. W., Raiffa, H., Schlaifer, R. y col. (1995). *Introduction to statistical decision theory*. MIT press. <https://mitpress.mit.edu/9780262662062/introduction-to-statistical-decision-theory/>

- Raiffa, H. (1970). *Decision Analysis Introductory Lectures on Choices Under Uncertainty. USA: Addison*. Wesley Publishing Company, Inc. <https://books.google.es/books?id=kS8zAAAAMAAJ>
- Ravinder, H. V. (1992). Bias in aggregations of subjective probability and utility. *Journal of the Operational Research Society*, 43(6), 621-627. <https://doi.org/10.1057/jors.1992.87>
- Renooij, S. & Witteman, C. (1999). Talking probabilities: communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning*, 22(3), 169-194. [https://doi.org/10.1016/S0888-613X\(99\)00027-4](https://doi.org/10.1016/S0888-613X(99)00027-4)
- Röver, C. (2017). Bayesian random-effects meta-analysis using the bayesmeta R package. *arXiv preprint arXiv:1711.08683*. <https://doi.org/10.18637/jss.v093.i06>
- Rust, R. T., Lehmann, D. R. & Farley, J. U. (1990). Estimating publication bias in meta-analysis. *Journal of Marketing Research*, 27(2), 220-226. <https://doi.org/10.1177/002224379002700209>
- Science, D. (& Laboratory), T. (2014). *The Probabilistic Elicitation of Subjective Data* (inf. téc.). DSTL/TR9234. United Kingdom. <https://www.gov.uk/government/publications/the-probabilistic-elicitation-of-subjective-data>
- Seo, J. I. & Kim, Y. (2022). Nonparametric prior elicitation for a binomial proportion. *Communications in Statistics-Simulation and Computation*, 51(6), 2809-2821. <https://doi.org/10.1080/03610918.2019.1702210>
- Sharp, S. J. & Thompson, S. G. (2000). Analysing the relationship between treatment effect and underlying risk in meta-analysis: comparison and development of approaches. *Statistics in medicine*, 19(23), 3251-3274. [https://doi.org/10.1002/1097-0258\(20001215\)19:23<3251::AID-SIM625>3.0.CO;2-2](https://doi.org/10.1002/1097-0258(20001215)19:23<3251::AID-SIM625>3.0.CO;2-2)
- Sinha, B. K., Hartung, J. & Knapp, G. (2011). *Statistical meta-analysis with applications*. John Wiley & Sons. <https://doi.org/10.1002/9780470386347>
- Sloman, S. A., Over, D., Slovak, L. & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational behavior and human decision processes*, 91(2), 296-309. [https://doi.org/10.1016/S0749-5978\(03\)00021-9](https://doi.org/10.1016/S0749-5978(03)00021-9)
- Slottje, P., Van der Sluijs, J., Knol, A. B. y col. (2008). *Expert Elicitation: Methodological suggestions for its use in environmental health impact assessments* (inf. téc.). National Institute for Public Health y the Environment. <https://rivm.openrepository.com/handle/10029/259123>
- Solomon, I. (1982). Probability assessment by individual auditors and audit teams: An empirical investigation. *Journal of Accounting Research*, 689-710. <https://doi.org/10.2307/2490893>
- Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of economic perspectives*, 15(3), 131-150. <https://doi.org/10.1257/jep.15.3.131>
- Strickler, Z. (1999). Elicitation methods in experimental design research. *Design issues*, 15(2), 27-39. <https://doi.org/10.2307/1511840>

- Swieringa, R., Gibbins, M., Larsson, L. & Sweeney, J. L. (1976). Experiments in the heuristics of human information processing. *Journal of Accounting Research*, *14*, 159-187. <https://doi.org/10.2307/2490450>
- Tangarife Quintero, J. (2016). *Elicitación de una distribución a priori para el modelo logístico* (Tesis doctoral). Facultad de Ciencias. <https://repositorio.unal.edu.co/handle/unal/57633>
- Team, R. C. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: a meta-analysis. *Journal of Risk Research*, *5*(2), 177-186. <https://doi.org/10.1080/13669870110038179>
- Truong, P. N. & Heuvelink, G. B. (2013). Uncertainty quantification of soil property maps with statistical expert elicitation. *Geoderma*, *202*, 142-152. <https://doi.org/10.1016/j.geoderma.2013.03.016>
- Tversky, A. (1974). Assessing uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 148-159. <https://www.jstor.org/stable/2984810>
- Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>
- van der Pas, J.-W. G., Marchau, V. A., Walker, W. E., Van Wee, G. & Vlassenroot, S. (2012). ISA implementation and uncertainty: A literature review and expert elicitation study. *Accident Analysis & Prevention*, *48*, 83-96. <https://doi.org/10.1016/j.aap.2010.11.021>
- Van Noordwijk, J. M., Dekker, A., Cooke, R. M. & Mazzuchi, T. A. (1992). Expert judgment in maintenance optimization. *IEEE Transactions on Reliability*, *41*(3), 427-432. <https://doi.org/10.1109/24.159813>
- van Dorp, R. é. & Kotz, S. (2002). A novel extension of the triangular distribution and its parameter estimation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *51*(1), 63-79. <https://doi.org/10.1111/1467-9884.00299>
- Villar, J., Mackey, M. E., Carroli, G. & Donner, A. (2001). Meta-analyses in systematic reviews of randomized controlled trials in perinatal medicine: comparison of fixed and random effects models. *Statistics in medicine*, *20*(23), 3635-3647. <https://doi.org/10.1002/sim.1096>
- Wallsten, T. & Budescu, D. (1980). Encoding subjective probabilities. *A psychological and psychometric review*. The University of North Carolina, Department of Psychology Chapel Hill: NC. <https://www.jstor.org/stable/2631269>
- Wilcox, C., Mallos, N. J., Leonard, G. H., Rodriguez, A. & Hardesty, B. D. (2016). Using expert elicitation to estimate the impacts of plastic pollution on marine wildlife. *Marine Policy*, *65*, 107-114. <https://doi.org/10.1016/j.marpol.2015.10.014>

- Winkler, R. L. (1967a). The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical Association*, 62(319), 776-800. <https://doi.org/10.2307/2283671>
- Winkler, R. L. (1967b). The quantification of judgment: Some methodological suggestions. *Journal of the American Statistical Association*, 62(320), 1105-1120. <https://doi.org/10.1080/01621459.1967.10500920>
- Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management science*, 15(2), B-61. <https://doi.org/10.1287/mnsc.15.2.B61>
- Winkler, R. L. (1969). Scoring rules and the evaluation of probability assessors. *Journal of the American Statistical Association*, 64(327), 1073-1078. <https://doi.org/10.1007/BF02562681>
- Wiser, R., Rand, J., Seel, J., Beiter, P., Baker, E., Lantz, E. & Gilman, P. (2021). Expert elicitation survey predicts 37% to 49% declines in wind energy costs by 2050. *Nature Energy*, 6(5), 555-565. <https://doi.org/10.1038/s41560-021-00810-z>
- Witteman, C. & Renooij, S. (2003). Evaluation of a verbal–numerical probability scale. *International Journal of Approximate Reasoning*, 33(2), 117-131. [https://doi.org/10.1016/S0888-613X\(02\)00151-2](https://doi.org/10.1016/S0888-613X(02)00151-2)

Juan Carlos Correa Morales

Doctor en Estadística

Profesor Asociado, Escuela de Estadística

Facultad de Ciencias

Universidad Nacional de Colombia, sede Medellín

Email: jccorrea@unal.edu.co

Carlos Javier Barrera Causil

Doctor en Ciencias Estadística

Profesor Titular 2, Facultad de Ciencias Exactas y Aplicadas

Institución Universitaria ITM

Medellín-Colombia

Email: carlosbarrera@itm.edu.co



Técnicas de elicitación de información personal para su uso en estadística bayesiana

Fuentes tipográficas: Computer Modern para texto corrido, en 12 puntos, para títulos en Computer Modern, en 24 puntos y subtítulos

El análisis bayesiano es una de las áreas de la estadística que más ha crecido en los últimos años, debido a las mejoras en la capacidad de cómputo y procesamiento de datos de los equipos actuales y a su capacidad para resolver problemas que no se pueden abordar con otros métodos. Además, los métodos bayesianos difieren de otros enfoques porque incorporan información previa en sus análisis. Dicha información puede provenir, entre otras fuentes, del conocimiento de expertos, lo que nos lleva a asumir importantes retos relacionados con la extracción y cuantificación de dicho conocimiento. En nuestro trabajo, presentamos diferentes formas de cuantificar información y expresarla a través de modelos de probabilidad, de manera que el lector disponga de metodologías empíricas claras que puedan usarse para procesos de estimación *a priori* o *a posteriori* en problemas de su interés. El texto comprende nueve capítulos que abordan desde definiciones básicas necesarias para el estudio y conocimiento de estas técnicas hasta propuestas metodológicas de elicitación y análisis de información de múltiples expertos. El contenido se ilustra con códigos desarrollados en el software R, con el fin de facilitar el estudio de la obra y que el lector pueda aplicar estas técnicas en sus investigaciones.

