Una primera aproximación a la Estadística Bayesiana. Sesión I

Asun Mtnez. Mayoral y Javier Morales

IUI CIO Universidad Miguel Hernández de Elche

Noviembre 2022

Objetivos de aprendizaje

- Conocer los conceptos básicos en el planteamiento bayesiano de la Estadística.
- Identificar la relevancia de la información previa y de expertos, y la proporcionada por los datos.
- Conocer los procedimientos básicos para conjugar la información disponible.
- Aplicar los conocimientos básicos en problemas sencillos.
- Descubrir las dificultades computacionales en la inferencia bayesiana.

Índice de contenidos propuesto

1. De probabilidad va la historia: probabilidades condicionadas y el teorema de Bayes.
2. La jerga base: incertidumbre, a priori, a posteriori y verosimilitud.
3. Manos en la masa 1: ¿con qué probabilidad ocurrió?
4. Manos en la masa 2: ¿con qué abundancia ocurrió?
5. Curioseando para saber más: manuales y software.

Índice de contenidos efectivo

1. SESIÓN 1: Probabilidades, Bayes y las proporciones.
2. SESIÓN 2: INLA y la regresión.
3. SESIÓN 3: INLA y el ANOVA.
4. SESIÓN 4: INLA y los GLM.

Probabilidad

Orígenes y evolución de la Estadística Bayesiana

Thomas Bayes (1702-1761), clérigo y científico-matemático investiga causas y efectos y propone la ley de probabilidad inversa en 1740s. Laplace descubre la regla de modo independiente en 1774. Durante los 150 años posteriores se impone sin embargo la estadística frecuentista de Fisher.

Orígenes y evolución de la Estadística Bayesiana

Thomas Bayes (1702-1761), clérigo y científico-matemático investiga causas y efectos y propone la ley de probabilidad inversa en 1740s. Laplace descubre la regla de modo independiente en 1774. Durante los 150 años posteriores se impone sin embargo la estadística frecuentista de Fisher.

Siglo XX:

- Turing, Kolmogorov, Jeffreys, Lindley, deFinetti, Savage, Bernardo, Berger, ... desarrollan la teoría Bayesiana.
- 1990-: Con el desarrollo de los ordenadores y algoritmia basada en simulación, se empiezan a plantear modos de resolver problemas.
- 1995-: Desarrollo de software: OpenBugs, WinBugs, librerías en R (Albert, 2009), INLA, 2019. Expansión notable en todos los ámbitos de conocimiento.

Evolución

- Hoy en día, los filtros de spam bayesianos nos protegen de correos fraudulentos.
- La regla de Bayes permite la localización de naúfragos supervivientes cuando se hunde un cayuco.
- Los científicos descubren cómo se controlan y regulan los genes. Rastrea la web y vende canciones y películas.
- Ha penetrado en la ciencia de la computación, la inteligencia artificial, el machine learning, Wall Street, astronomía y física, seguridad del hogar, Microsoft y Google.
- Ayuda a los computadores a traducir de un lenguaje a otro.
- Se ha convertido en una metáfora para explicar cómo aprende y funciona nuestro cerebro.
- Se generan recomendaciones bayesianas a las agencias gubernamentales sobre energía, educación e investigación.

(Fte: McGraine, 2011)

Principios (I)

- El modo bayesiano de razonar es el modo habitual de razonar.
  - ¿Qué es un fenómeno extraordinario?
  - ¿Qué tiempo hará?
  - ¿Cuál es el origen de ese sujeto?
  - ¿Me voy a curar de una enfermedad?
  - ¿Me va a subir el colesterol?

Principios (I)

- El modo bayesiano de razonar es el modo habitual de razonar.
    - ¿Qué es un fenómeno extraordinario?
    - ¿Qué tiempo hará?
    - ¿Cuál es el origen de ese sujeto?
    - ¿Me voy a curar de una enfermedad?
    - ¿Me va a subir el colesterol?

¿Cómo analizamos la evidencia, cambiamos nuestro pensamiento y conseguimos nueva información, para tomar decisiones racionales en función de la incertidumbre?

Principios (I)

- El modo bayesiano de razonar es el modo habitual de razonar.
  - ¿Qué es un fenómeno extraordinario?
  - ¿Qué tiempo hará?
  - ¿Cuál es el origen de ese sujeto?
  - ¿Me voy a curar de una enfermedad?
  - ¿Me va a subir el colesterol?

¿Cómo analizamos la evidencia, cambiamos nuestro pensamiento y conseguimos nueva información, para tomar decisiones racionales en función de la incertidumbre?

Actualizamos nuestra información inicial sobre algo cuando recabamos nueva información objetiva con la que mejoramos nuestro conocimiento.

Principios (II)

- No se acude al muestreo repetido.
  - ¿Cuál es la probabilidad de que un volcán erupcione?
  - ¿Cuál es la probabilidad de sobrevivir con un tratamiento médico que se actualiza con nuevos fármacos?

Principios (II)

- No se acude al muestreo repetido.
  - ¿Cuál es la probabilidad de que un volcán erupcione?
  - ¿Cuál es la probabilidad de sobrevivir con un tratamiento médico que se actualiza con nuevos fármacos?
- La inferencia no se basa en estimadores y en sus propiedades en el muestreo, sino en el conocimiento adquirido sobre los parámetros de interés.

Principios (II)

- No se acude al muestreo repetido.
  - ¿Cuál es la probabilidad de que un volcán erupcione?
  - ¿Cuál es la probabilidad de sobrevivir con un tratamiento médico que se actualiza con nuevos fármacos?
- La inferencia no se basa en estimadores y en sus propiedades en el muestreo, sino en el conocimiento adquirido sobre los parámetros de interés.
- El teorema de Bayes permite actualizar información, sea cual sea su fuente (mínima, expertos, datos, estudios previos, ...).

El teorema de Bayes (I)

- Probabilidad marginal de un evento. ¿Con qué probabilidad sucede A en una serie de eventos?

$$\text{Probabilidad(A)} = \frac{\text{Casos en que sucede A}}{\text{Casos totales}}$$

El teorema de Bayes (I)

- Probabilidad marginal de un evento. ¿Con qué probabilidad sucede A en una serie de eventos?

$$\text{Probabilidad(A)} = \frac{\text{Casos en que sucede A}}{\text{Casos totales}}$$

- Probabilidad condicionada. ¿Con qué probabilidad sucede A si sabemos que ha sucedido B?

$$\text{Probabilidad (A|B)} = \frac{\text{Casos en que suceden A y B}}{\text{Casos en que sucede B}}$$

- Probabilidad conjunta. ¿Con qué probabilidad suceden A y B?

El teorema de Bayes (II)

El **teorema de Bayes** relaciona probabilidades marginales, conjuntas y condicionadas de dos sucesos A y B:

$$Pr(A, B) = Pr(A|B) \cdot Pr(B) = Pr(B|A) \cdot Pr(A)$$

$$Pr(A|B) = \frac{Pr(A, B)}{Pr(B)} = \frac{Pr(B|A) \cdot Pr(A)}{Pr(B)}$$

Y recordamos el teorema de probabilidad total:

$$Pr(A) = \sum_{B} Pr(A|B) Pr(B)$$

Calculamos probabilidades (I)

Estamos interesados en calcular la probabilidad de ir al cielo si eres bayesiano. Nos dan algo de información:

- 6 personas de cada 9 que van al infierno son bayesianos
- 5 personas de cada 7 que van al cielo son bayesianos
- el 75% de la población va al infierno
- el 25% de la población va al cielo.

(Fte: LaplacesDemon/BayesianInference)

Calculamos probabilidades (II)

Reescribimos toda la información en términos de probabilidades:

- ¿$Pr(cielo|Bayesian)$?
- $Pr(Bayes|infierno) = 6/9$
- $Pr(Bayes|cielo) = 5/7$
- $Pr(infierno) = 0.75$
- $Pr(cielo) = 0.25$

Calculamos probabilidades (II)

Reescribimos toda la información en términos de probabilidades:

- ¿$Pr(cielo|Bayesian)$?
- $Pr(Bayes|infierno) = 6/9$
- $Pr(Bayes|cielo) = 5/7$
- $Pr(infierno) = 0.75$
- $Pr(cielo) = 0.25$

$$
\begin{aligned}
Pr(cielo|Bayes) &= \frac{Pr(Bayes|infierno)Pr(infierno)}{Pr(Bayes)} \\
&= \frac{Pr(Bayes|infierno)Pr(infierno)}{Pr(Bayes|cielo)Pr(cielo) + Pr(Bayes|infierno)Pr(infierno)} \\
&= \frac{(6/9)(0.75)}{(5/7)(0.25) + (6/9)(0.75)} = \frac{0.5}{0.6786} = 0.74
\end{aligned}
$$

Conceptos

Conceptos (I)

- Un **parámetro** es una característica poblacional de interés, desconocida, sobre la que pretendemos inferir a partir de datos observados.

1. La probabilidad de que una planta se infecte por un hongo.
2. El valor medio de un índice de contaminación medioambiental en una ubicación.

Conceptos (I)

- Un **parámetro** es una característica poblacional de interés, desconocida, sobre la que pretendemos inferir a partir de datos observados.

1. La probabilidad de que una planta se infecte por un hongo.
2. El valor medio de un índice de contaminación medioambiental en una ubicación.

- Las **variables** identifican las características que hemos de observar en la población de interés para obtener información sobre el parámetro de interés.

1. Una planta tiene el hongo: sí/no.
2. Valor del índice.

Conceptos (II)

- Los **datos** con que inferimos provienen de observaciones de la característica de interés sobre un sector/muestra de la población de interés.

1. Observación y registro de una selección de plantas en una finca.
2. Observación y registro del índice a lo largo del tiempo.

Conceptos (II)

- Los **datos** con que inferimos provienen de observaciones de la característica de interés sobre un sector/muestra de la población de interés.

1. Observación y registro de una selección de plantas en una finca.
2. Observación y registro del índice a lo largo del tiempo.

- La **verosimilitud** proviene de la distribución asumida para la variable observada, usando todos los datos disponibles. Informa sobre qué valores del parámetro son más/menos verosímiles a la vista de los datos observados.

1. La distribución binomial con probabilidad desconocida.
2. El producto de densidades normales para los datos observados, con media desconocida.

Conceptos (III)

En Estadística Clásica (frecuentista), son importantes:

- Los **estimadores**: funciones de los datos cuya distribución en el muestreo se obtiene de la distribución asumida sobre la variable observada. Utilizada para resolver las inferencias sobre el parámetro de interés en términos de probabilidad, valor esperado y error estándar.

Conceptos (III)

En Estadística Clásica (frecuentista), son importantes:

- Los **estimadores**: funciones de los datos cuya distribución en el muestreo se obtiene de la distribución asumida sobre la variable observada. Utilizada para resolver las inferencias sobre el parámetro de interés en términos de probabilidad, valor esperado y error estándar.
- Las **estimaciones**: valores que toman los estimadores para unos datos observados. El error estándar del estimador calculado con los datos observados da el error estándar de la estimación.

Conceptos (III)

En Estadística Clásica (frecuentista), son importantes:

- Los **estimadores**: funciones de los datos cuya distribución en el muestreo se obtiene de la distribución asumida sobre la variable observada. Utilizada para resolver las inferencias sobre el parámetro de interés en términos de probabilidad, valor esperado y error estándar.
- Las **estimaciones**: valores que toman los estimadores para unos datos observados. El error estándar del estimador calculado con los datos observados da el error estándar de la estimación.
- Un **contraste de hipótesis** plantea en la hipótesis alternativa lo que queremos demostrar sobre el parámetro de interés, y en la nula el suceso complementario.

Conceptos (III)

En Estadística Clásica (frecuentista), son importantes:

- Los **estimadores**: funciones de los datos cuya distribución en el muestreo se obtiene de la distribución asumida sobre la variable observada. Utilizada para resolver las inferencias sobre el parámetro de interés en términos de probabilidad, valor esperado y error estándar.
- Las **estimaciones**: valores que toman los estimadores para unos datos observados. El error estándar del estimador calculado con los datos observados da el error estándar de la estimación.
- Un **contraste de hipótesis** plantea en la hipótesis alternativa lo que queremos demostrar sobre el parámetro de interés, y en la nula el suceso complementario.
- Los problemas de estimación se resuelven en términos de estimaciones, errores estándar e intervalos de confianza. Los problemas de contraste de hipótesis en términos de p-valores.

Inferir en clásico

Inferencia clásica (I)

Se quiere testar si $p = Pr(\text{Mejoría})$ con el fármaco Z es superior a 0.8, esto es,

$$\begin{aligned} H_0 &: \quad p \leq 0.8 \\ H_1 &: \quad p > 0.8. \end{aligned}$$

Inferencia clásica (I)

Se quiere testar si $p = Pr(\text{Mejoría})$ con el fármaco Z es superior a 0.8, esto es,

$$\begin{aligned} H_0 &: \quad p \leq 0.8 \\ H_1 &: \quad p > 0.8. \end{aligned}$$

Observamos n pacientes y registramos 1/0 (mejora/no).

$$X_i \sim Br(p) \rightarrow Y_n = \sum_{i=1}^{n} X_i \sim Bin(n, p)$$

Estadístico de contraste:

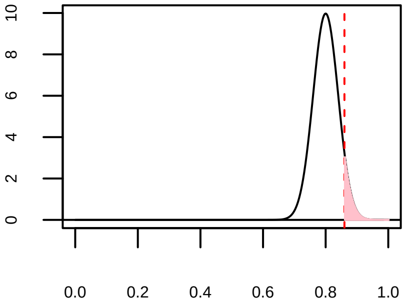
$$\hat{p}_n = \frac{Y_n}{n} \sim N(p, p(1-p)/n)$$

Inferencia clásica (II)

Distribución del estadístico de contraste bajo H_0:

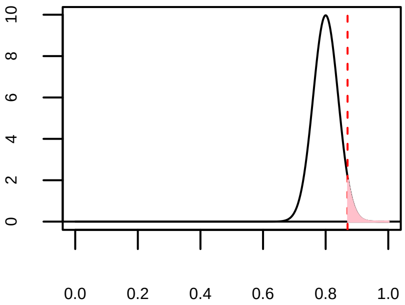
$$\hat{p}_n | H_0 \sim N(0.8, 0.8(0.2)/n)$$

- Observados $y =$ **86 en** $n =$ **100**.
- Estimación e IC(95%) para p : 0.86(0.79,1).
- P-valor $= Pr(\hat{p}_n > y_n/n | H_0) = 0.08$. ¿Conclusión?



Inferencia clásica (III)

- Observados $y = \mathbf{87}$ **en** $n = \mathbf{100}$.
- Estimación e IC(95%) para p : 0.87(0.801,1).
- P-valor $= Pr(\hat{p}_n > y_n/n|H_0) = 0.047$. ¿Conclusión?



Inferencia clásica (IV)

Desde el enfoque frecuentista podemos calcular ciertas probabilidades como:

- ¿Cuál es la probabilidad de que el EMV tome un valor superior a la estimación obtenida cuando la hipótesis nula es cierta?
- ¿Cuál es la probabilidad de rechazar H_0 cuando es cierta?
- ¿Cuáles son las evidencias a favor de la hipótesis de interés (H_1)?
- ¿Qué porcentaje de veces (en el muestreo repetido) el intervalo de confianza "pescará" al verdadero valor de p, cuando la hipótesis nula es cierta?
- ¿Qué probabilidad tenemos de equivocarnos rechazando H_0 (aceptando H_1)?

Inferencia clásica (V)

... si bien no otras que nos interesan:

- ¿Qué certidumbre tenemos sobre $H_1 : p > 0.8$? El p-valor NO es la probabilidad de $p \leq 0.8$.
- ¿Qué valor esperamos para p y con qué certidumbre?
- ¿Con qué probabilidad mejorará un nuevo enfermo que aparezca en el estudio?
- Si se decide repetir el experimento con otros 50 enfermos, ¿cuál es la probabilidad de que mejoren más de 40?

Inferir en Bayesiano

Inferir en Bayesiano. Probabilidad

Razonamos en términos de PROBABILIDAD

- sobre lo observable y variable en el muestreo: datos
- y también sobre lo desconocido e inobservable: parámetros

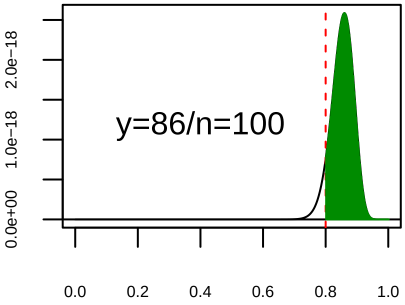
INCERTIDUMBRE $\equiv$ PROBABILIDAD

Lo que sabemos (y no sabemos) lo expresamos con distribuciones de probabilidad.

Inferir en Bayesiano. Verosimilitud

Verosimilitud Es la información que proporcionan los datos sobre los parámetros de interés usando la distribución f asumida para modelizarlos, $l(\theta|x) \propto f(x|\theta)$.

$$Y_n \sim Bin(n,p) \rightarrow l(p|y_n) \propto p^{y_n} \cdot (1-p)^{n-y_n}$$



Inferir en Bayesiano. Distribución a priori

Los método bayesianos requieren una distribución a priori sobre TODOS los parámetros desconocidos. ¿Cómo conseguirla?

- **Información a priori objetiva**. A priori no se sabe nada sobre el parámetro de interés. Un análisis objetivo se espera que proporcione resultados tan objetivos como los producidos por un análisis frecuentista.

Inferir en Bayesiano. Distribución a priori

Los método bayesianos requieren una distribución a priori sobre TODOS los parámetros desconocidos. ¿Cómo conseguirla?

- **Información a priori objetiva**. A priori no se sabe nada sobre el parámetro de interés. Un análisis objetivo se espera que proporcione resultados tan objetivos como los producidos por un análisis frecuentista.

- **Información a priori subjetiva**, especificada en términos de una distribución de probabilidad obtenida a partir de:
  - Información de expertos en el área.
  - Información de experimentos previos.

Inferir en Bayesiano. Distribución a priori

Distribución de probabilidad para expresar el conocimiento a priori, antes de observar los datos, sobre los parámetros, $\pi(\theta)$.

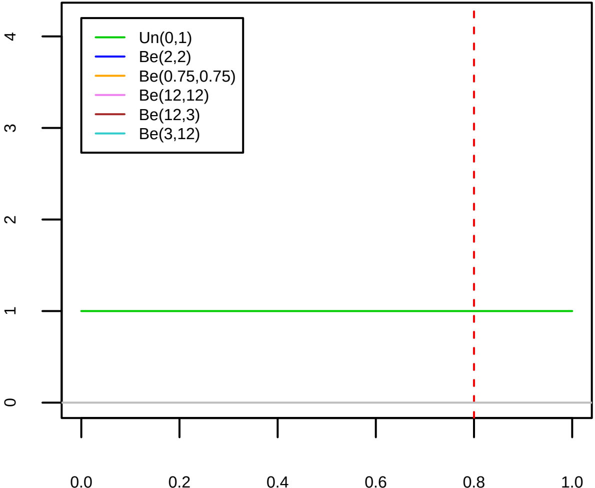
- Mínima información: $p \sim Un(0,1)$
- Información a priori: $Be(\alpha, \beta) \equiv Be(\alpha, \alpha k)$ en términos de media m y varianza v

$$k = \frac{1}{m} - 1; \alpha = \frac{(1 - m - v/m)m^2}{v}, \text{ para } v < m(1-m).$$

Nota: $Un(0,1) \equiv Be(1,1)$

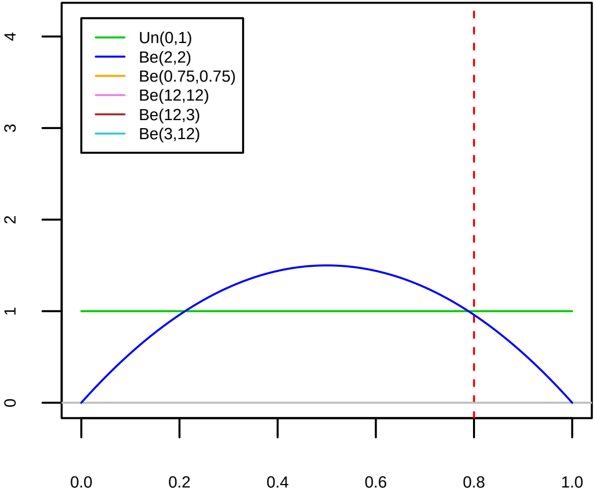
Inferir en Bayesiano. Distribución a priori

Distribución a priori difusa o mínimo-informativa
(Media=0.5,Var=0.08)



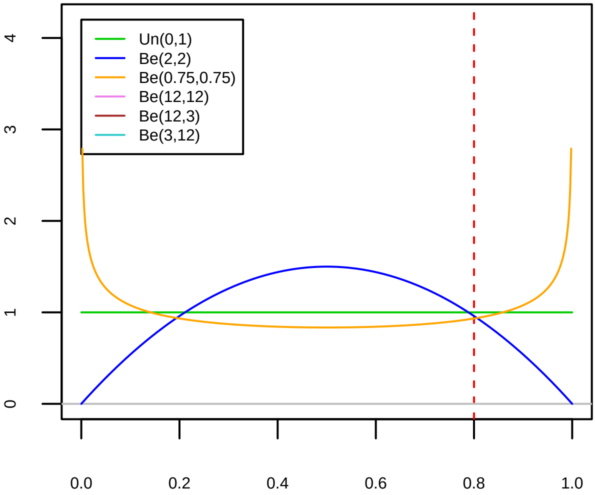
Inferir en Bayesiano. Distribución a priori

Distribución a priori $Be(2,2)$ (Media=0.5,Var=0.05)



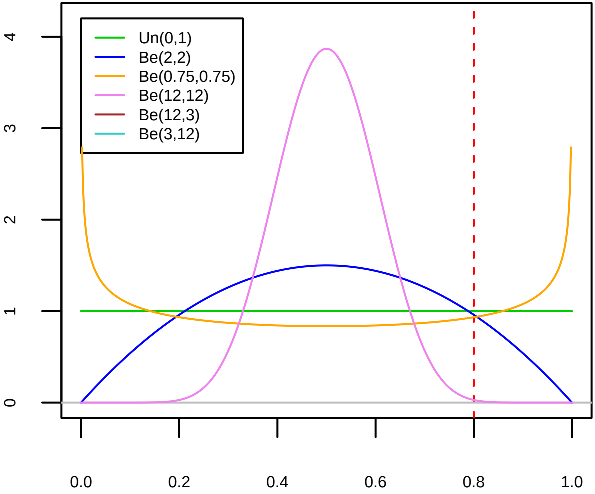
Inferir en Bayesiano. Distribución a priori

Distribución a priori $Be(0.75, 0.75)$ (Media=0.5,Var=0.1)



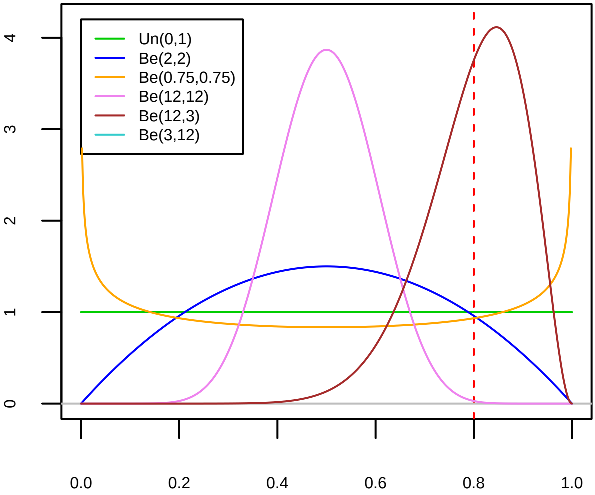
Inferir en Bayesiano. Distribución a priori

Distribución a priori $Be(12, 12)$ (Media=0.5,Var=0.01)



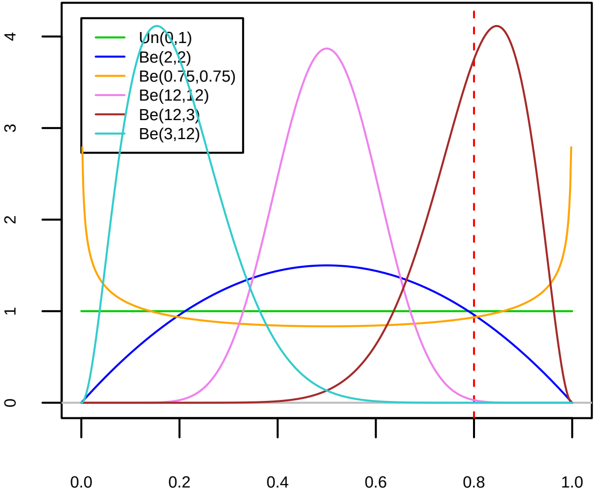
Inferir en Bayesiano. Distribución a priori

Distribución a priori $Be(12, 3)$ (Media=0.8,Var=0.01)



Inferir en Bayesiano. Distribución a priori

Distribución a priori $Be(3, 12)$ (Media=0.2,Var=0.01)



Inferir en Bayesiano. Distribución posterior

Las inferencias sobre los parámetros de interés se resuelven con su **distribución posterior**. Esta surge al aplicar el teorema de Bayes y constituye un compromiso entre la Información inicial en la distribución a priori y la verosimilitud o modelo asumido para los datos. Proporciona:

- Estimación puntual: media, mediana, moda,etc. . .
- Precisión de la estimación: varianza, desviación típica
- Estimación en intervalo: Región Creíble (RC), o región con máxima densidad posterior (HPDR).
- Probabilidad de cualquier hipótesis sobre los parámetros: . . . modelización específica para hipótesis puntuales.

Inferir en Bayesiano. Distribución posterior

Actualizamos información combinando la información de los datos y la información previa con el Teorema de Bayes.

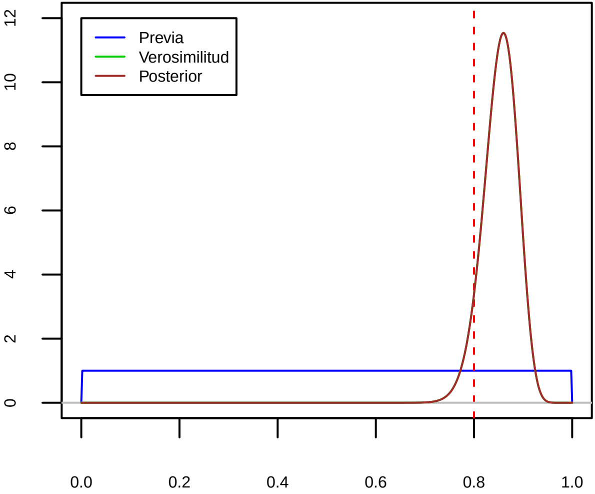
$$\pi(\theta|x) = \frac{f(x,\theta)\pi(\theta)}{f(x)} = \frac{f(x|\theta)\pi(\theta)}{\int_\theta f(x|\theta)} = \frac{l(\theta|x)\pi(\theta)}{\int_\theta l(\theta|x)\pi(\theta)}$$

Para nuestra $p = Pr(\text{mejoría})$:

$$\begin{aligned} \pi(p|y_n) &\propto [p^{\alpha-1}(1-p)^{\beta-1}]\cdot[p^{y_n}\cdot(1-p)^{n-y_n}] \\ &\propto p^{\alpha-1+y_n}(1-p)^{\beta-1+n-y_n} \\ &\propto Be(p|\alpha+y_n,\beta+n-y_n) \end{aligned}$$

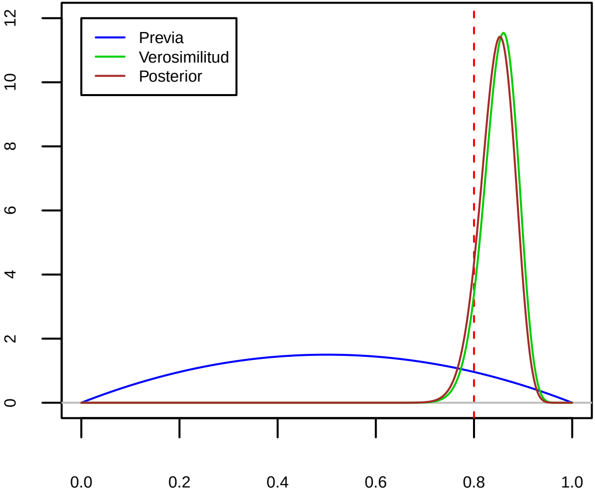
Inferir en Bayesiano. Distribución posterior

Distribución posterior con previa mín.info $Un(0,1)$



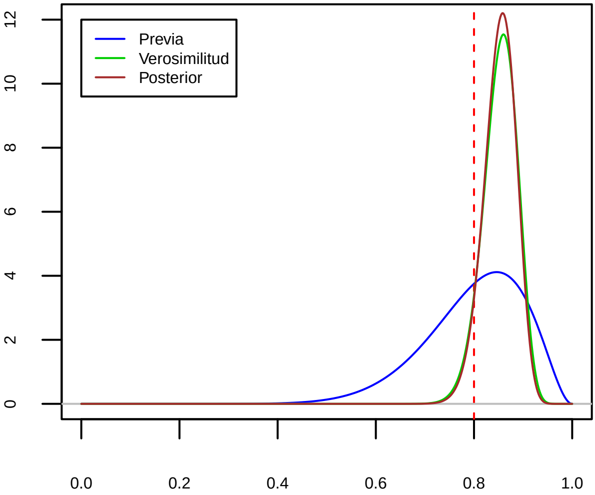
Inferir en Bayesiano. Distribución posterior

Distribución posterior con previa dispersa $Be(2, 2)$



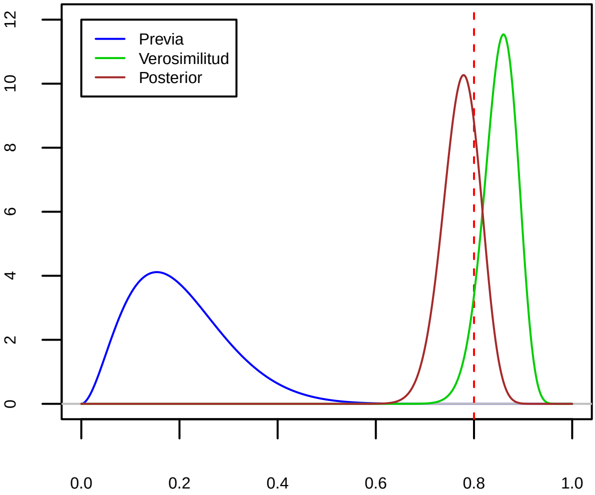
Inferir en Bayesiano. Distribución posterior

Distribución posterior con previa informativa $Be(12, 3)$



Inferir en Bayesiano. Distribución posterior

Distribución posterior con previa sesgada $Be(3, 12)$



Inferir en Bayesiano. Respuestas estimación

Podemos responder las cuestiones que nos interesan:

- ¿Qué valor esperamos para p y con qué certidumbre?

Media posterior:

$$E(p|y_n) = E[Be(\alpha + y_n, \beta + n - y_n)] = \frac{\alpha + y_n}{\alpha + \beta + n}$$

Varianza posterior:

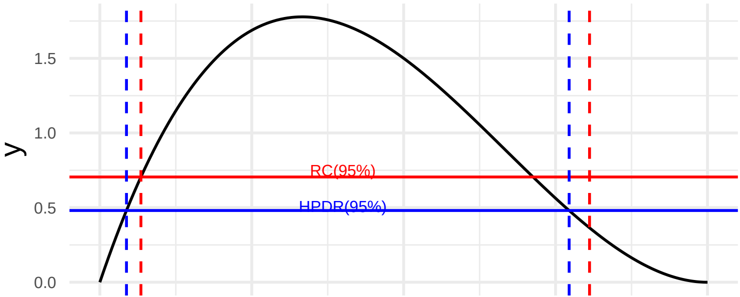
$$Var(p|y_n) = Var[Be(\alpha + y_n, \beta + n - y_n)] = \frac{(\alpha + y_n)(\beta + n - y_n)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)}$$

Inferir en Bayesiano. Estimación en intervalo

Regiones creíbles y HPDR *(highest posterior density region)*:

$$RC(1-\alpha) = [q_{\alpha/2}, q_{1-\alpha/2}] \leftarrow\text{--}\ \text{cuantiles D.Posterior}$$
$$RC(95\%) = [q_{2.5\%}, q_{97.5\%}]$$
$$HPD(1-\alpha) = [a, b] \leftarrow\text{--}\ \text{con } \pi(a|datos) = \pi(b|datos)$$

y siempre $Pr[\theta \in RC(1-\alpha)|datos] = Pr[\theta \in HPD(1-\alpha)|datos] = 1-\alpha$



Inferir en Bayesiano. Respuestas estimación

Observados: $y_n = 86$, $n = 100$.

	E	sd	RC.low	RC.up	HPD.low	HPD.up
Un(0,1)	0.853	0.035	0.778	0.914	0.784	0.918
Be(2,2)	0.846	0.035	0.771	0.909	0.776	0.912
Be(12,3)	0.852	0.033	0.782	0.911	0.787	0.914
Be(3,12)	0.774	0.039	0.694	0.845	0.697	0.848

RETO: Obtén respuestas cuando varía el número de pacientes con mejoría y_n.

Inferir en Bayesiano. Contrastes de hipótesis

Plantear un contraste de hipótesis conlleva asumir supuestos a priori coherentes con el test, de forma que siempre es posible evaluar las **evidencias a priori** a favor de H_0 (y en contra de H_1), esto es,

$$Pr(H_0)/Pr(H_1).$$

Tras actualizar información, el contraste se resuelve en términos de:

- **probabilidades a posteriori**:

$$\text{Aceptar } H_0 \text{ si } Pr(H_0 \mid datos) > Pr(H_1 \mid datos)$$

- **Factor Bayes a favor de** H_0, o evidencias a favor de H_0 (y en contra de H_1), aportadas por los datos:

$$B_{01} = \frac{Pr(H_0 \mid datos)/Pr(H_1 \mid datos)}{Pr(H_0)/Pr(H_1)} = \frac{Pr(H_0 \mid datos)Pr(H_1)}{Pr(H_1 \mid datos)Pr(H_0)}$$

Se acepta H_0 si B_{01} resulta suficientemente grande.

Inferir en Bayesiano. Respuestas contraste

Podemos responder las cuestiones que nos interesan:

- ¿Qué certidumbre tenemos sobre $H_1 : p > 0.8$?

Probabilidad posterior:

$$Pr(H_1|y_n) = 1 - F_{Be(\alpha+y_n,\beta+n-y_n)}(0.8)$$

	PH1.ini	PH1.post	B01
Un(0,1)	0.200	0.926	0.020
Be(2,2)	0.104	0.899	0.013
Be(12,3)	0.552	0.934	0.087
Be(3,12)	0.000	0.259	0.000

Inferir en Bayesiano. Predicción

- ¿Con qué probabilidad mejorará un nuevo enfermo que aparezca en el estudio?

$$Pr(Z_1 = 1 | y_n = 86, n = 100)$$

- Si se decide repetir el experimento con otros 50 enfermos, ¿cuál es la probabilidad de que mejoren más de 40?

$$E(Z_{50} | y_n = 86, n = 100);\; RC(Z_{50} | y_n = 86, n = 100)$$

$$Pr(Z_{50} > 40 | y_n = 86, n = 100)$$

Distribución predictiva. Información actualizada sobre la variable observable $Z = Y$ conociendo una respuesta de un estudio previo $Y = y$.

$$f(z|y) = \int f(z|\theta, y)\pi(\theta|y)d\theta = \int f(z|\theta)\pi(\theta|y)d\theta$$

Inferir en Bayesiano. Predicción con simulación

Para predecir el nº de mejorías Z_m en m pacientes nuevos, conociendo y_n, podemos aproximar con simulación Monte Carlo:

$$p^i \sim \pi(p|y_n) \rightarrow z_m^i \sim f(z_m|p^i), \ i = 1, ..., M$$

Para inferir con ellos

$$E(Z_m|y_n) \approx \sum_1^M z_m^i / M$$

$$sd(Z_m|y_n) \approx sd\{z_m^i, i = 1, ..., M\}$$

$$RC95(Z_m) \approx [q_{0.025}, q_{0.975}]_{\{z_m^i, i=1,...,M\}}$$

Inferir en Bayesiano. Respuestas predicción

Observados: $y_n = 86$, $n = 100$.

- ¿Con qué probabilidad mejorará un nuevo enfermo que aparezca en el estudio? 1
- Si se decide repetir el experimento con otros 50 enfermos,
  - ¿cuántos se espera que mejoren? 42.6764
  - ¿con qué certidumbre? IC(95%)=[36,48]
  - ¿cuál es la probabilidad de que mejoren más de 40? 0.769

RETO. Calcula las respuestas con otros resultados en el estudio previo.

Conclusiones

- Utilización de muchos parámetros y estructuración jerárquica de los modelos.
- Chequeo de modelos: bondad de ajuste y comparación de predicciones con la realidad.
- Ir más allá de simples estimaciones puntuales. Uso de distribuciones completas.
- Uso de modelos probabilísticos para entender la realidad.
- Uso de simulación y computación.
- Inclusión de toda la información disponible posible.
- Inferencias robustas a las asunciones sobre los modelos.

Comparación de proporciones

Comparación de proporciones

Se está experimentando una nueva técnica quirúrgica en enfermos cardíacos para reducir el número de muertos durante el proceso post-operatorio. Se dispone de información proporcionada por:

- Hospital A con técnica antigua: 31 muertos en 215 operaciones
- Hospital B con técnica nueva: 29 muertos en 256 operaciones

Se desea:

- Estudiar la tasa de mortalidad en cada hospital.
- Comparar las tasas de ambos hospitales.

Comparación de proporciones. Verosimilitud

$Y_j =$ nº exitus de n_j operaciones, $p_j =$ probabilidad de exitus en el hospital j.

Modelización de las respuestas en cada hospital:

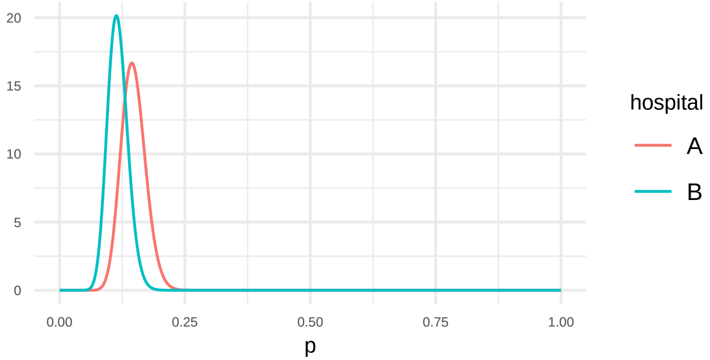
$$Y_j | p_j, n_j \sim Bin(n_j, p_j); \quad j = A, B$$

Verosimilitud conjunta (hospitales independientes/intercambiables):

$$l(p_A, p_B | y_A, y_B) \propto \prod_{j \in \{A,B\}} p_j^{y_j} (1 - p_j)^{n_j - y_j}$$

Comparación de proporciones. Verosimilitud

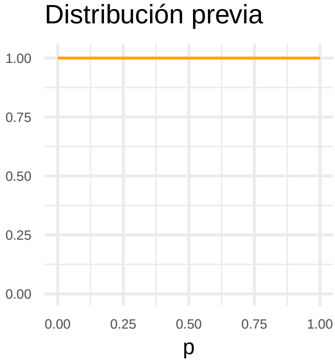
Verosimilitud en cada hospital



Comparación de proporciones. Información inicial objetiva

(NoInf) Ignorancia o mínima-información: p_j es una probabilidad

$$p \in [0,1] \Rightarrow p \sim Un(0,1) \equiv Be(1,1)$$



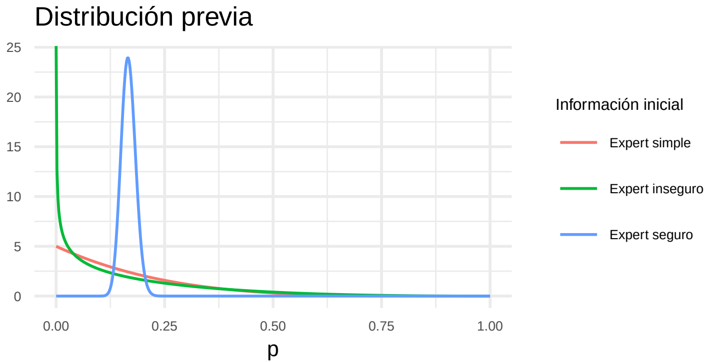
Comparación de proporciones. Información inicial experta

(Expert) Información experta: Cirujanos cardiovasculares estiman n_e muertes por cada n_o operaciones (media $m = n_e/n_o$), y manifiestan su incertidumbre con la desviación típica v.

$$p \sim Be(\alpha, \alpha\kappa);\ \kappa = \frac{1}{m} - 1;\ \alpha = \frac{m^2}{v^2}(1 - m - \frac{v^2}{m})$$

- Info: por cada 6 operaciones fallece un paciente.
- Certidumbre:
  - Expert simple: $Beta(\#\text{éxitos}, \#\text{fracasos}) \equiv Be(1, 5)$.
  - Expert inseguro: $sd(p) = E(p) \Rightarrow p \sim Be(0.67, 3.33)$
  - Expert seguro: $sd(p) = 10\%E(p) \Rightarrow p \sim Be(83.17, 415.33)$

Comparación de proporciones. Información inicial experta

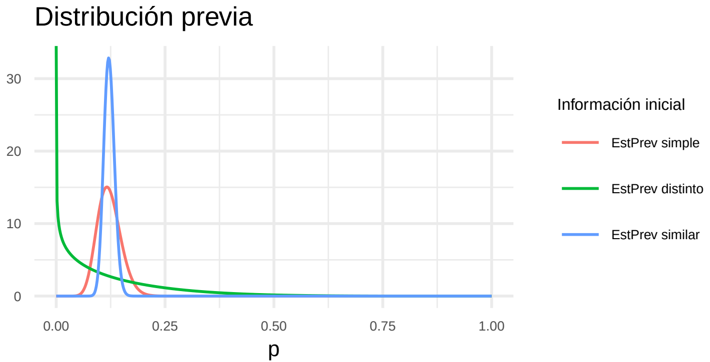


Comparación de proporciones. Información inicial de estudios previos

(EstPrev) Estudios previos más/menos similares en objetivos/población.

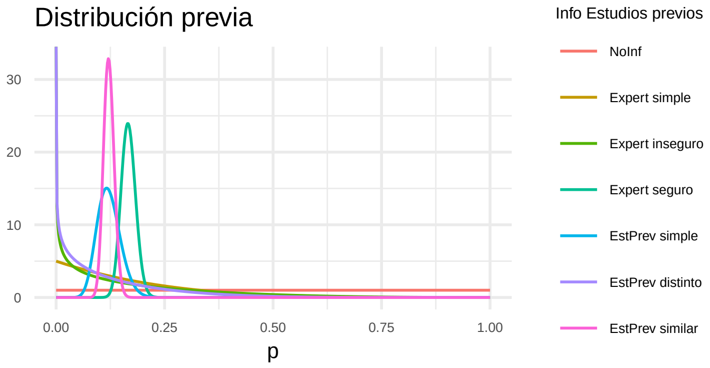
- Info: En otro estudio resultaron 18 muertes en 148 operaciones
- Certidumbre:
  - EstPrev simple: $Beta(\#\text{éxitos}, \#\text{fracasos}) \equiv Be(18, 130)$.
  - EstPrev distinto: $sd(p) = E(p) \Rightarrow p \sim Be(0.76, 5.47)$
  - EstPrev similar: $sd(p) = 10\% E(p) \Rightarrow p \sim Be(87.72, 633.51)$

Comparación de proporciones. Información inicial de estudios previos



Comparación de proporciones. Información inicial

Información inicial con todas las propuestas.



Comparación de proporciones. Inferencias y distribución posterior

Inferencias con la **distribución posterior**, obtenida de la a priori $\pi(\theta)$ y la verosimilitud $l(\theta|x) \propto f(x|\theta)$, donde $f(x|\theta)$ es el modelo probabilístico asumido para los datos a observar en la población caracterizada por el parámetro θ.

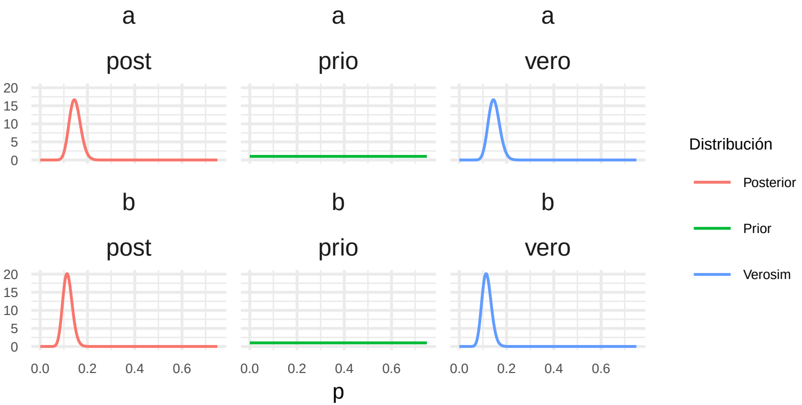
$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)} \propto l(\theta|x)\pi(\theta)$$

En nuestro caso, asumiendo previas similares para p_A y p_B:

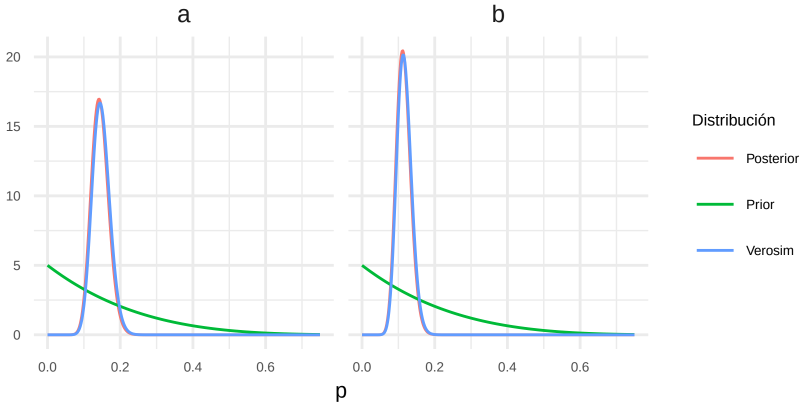
$$\begin{aligned} \pi(p_A, p_B|y_A, y_B) &\propto \prod_{j\in\{A,B\}} p_j^{y_j}(1-p_j)^{n_j-y_j}\ p_j^{\alpha_j-1}(1-p_j)^{\alpha_j\kappa_j-1} \\ &\propto Be(p_A|\alpha+y_A, \alpha\kappa+n_A-y_A)\cdot \\ &\quad \cdot Be(p_B|\alpha+y_B, \alpha\kappa+n_B-y_B) \end{aligned}$$

Comparación de proporciones. Posterior con previa NoInf

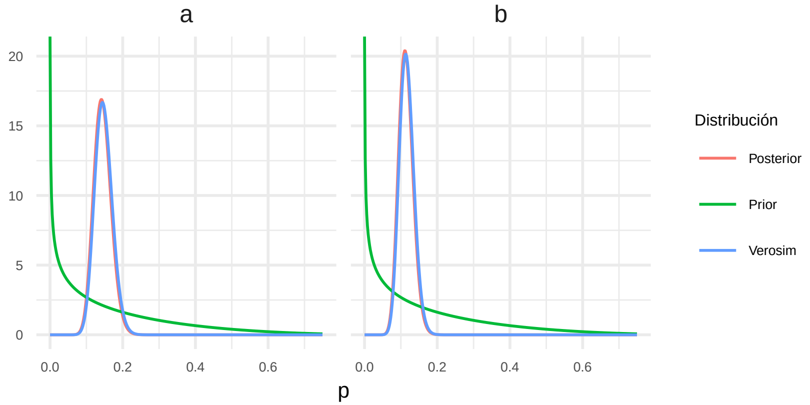
Los datos guían la posterior:



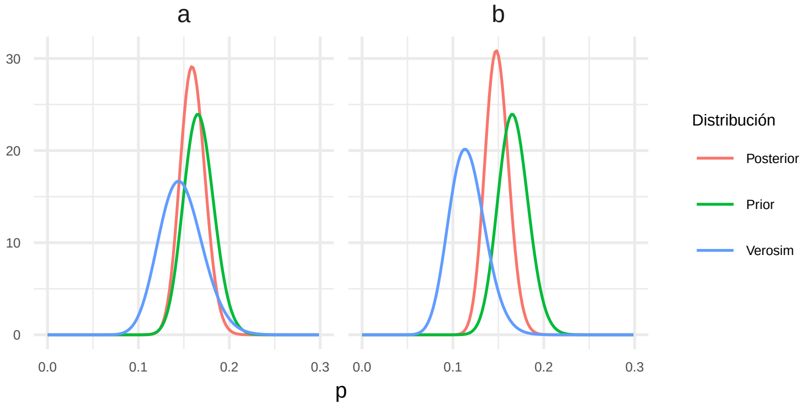
Comparación de proporciones. Posterior con previa Expert sencilla



Comparación de proporciones. Posterior con previa Expert inseguro

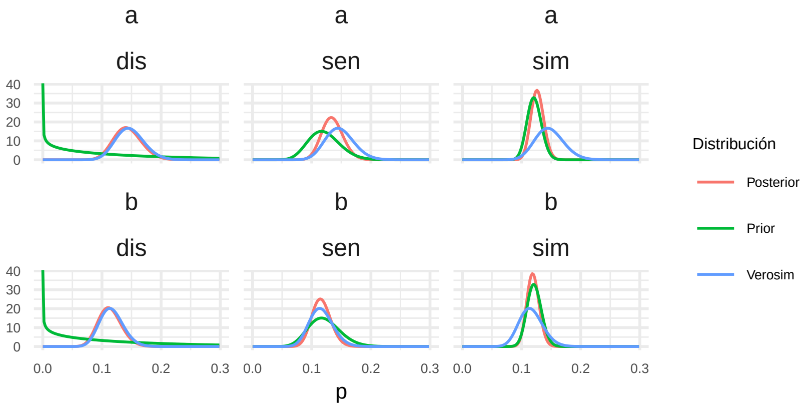


Comparación de proporciones. Posterior con previa Expert seguro



Comparación de proporciones. Posterior con Estudios Previos (EstPrev)

Información previa sencilla (sen), estudios similares (sim), estudios distintos (dis).



Comparación de proporciones. Inferencias sobre p_A

A priori	Media	sd	RC 0.95	HPD 0.95
NoInf	0.148	0.024	(0.104, 0.198)	(0.102, 0.195)
Expert Sencillo	0.145	0.024	(0.102, 0.194)	(0.100, 0.192)
Expert Seguro	0.160	0.014	(0.134, 0.188)	(0.133, 0.187)
Expert Inseguro	0.145	0.024	(0.101, 0.194)	(0.100, 0.192)
EstPrev Sencillo	0.135	0.018	(0.102, 0.172)	(0.101, 0.171)
EstPrev Similar	0.127	0.011	(0.106, 0.149)	(0.106, 0.148)
EstPrev Distinto	0.144	0.024	(0.101, 0.193)	(0.099, 0.190)

Comparación de proporciones. Inferencias sobre p_A y p_B

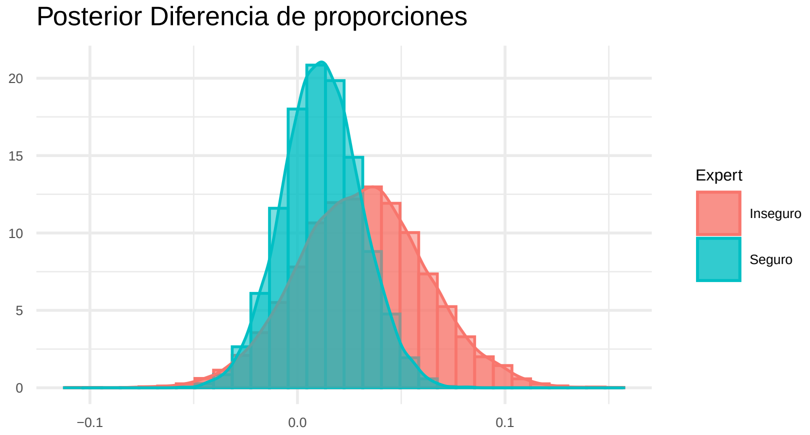
Pero estamos interesados en comparar p_A y p_B, de modo que necesitamos la distribución **posterior conjunta**, que asumiendo independencia entre los hospitales es:

$$\pi(p_A, p_B \mid y_A, y_B) = \pi(p_A \mid y_A)\pi(p_B \mid y_B).$$

Para obtener la distribución posterior de $p_A - p_B$ "de un modo fácil", basta con simular M valores de las correspondientes distribuciones posteriores $\pi(p_A \mid y_A)$ y $\pi(p_B \mid y_B)$, para con ellas obtener las estimaciones MonteCarlo que nos interesen:

$$\{p_A^{(i)}, p_B^{(i)}\}_{i=1}^M \Rightarrow \{p_A^{(i)} - p_B^{(i)}\}_{i=1}^M \sim \pi(p_A - p_B \mid y_A, y_B)$$

Comparación de proporciones. Resultados Expert



Comparación de proporciones. Resultados Expert

Estimación posterior de la diferencia de proporciones

Expert	m	s	rc.low	rc.up	hpd.low	hpd.up
Inseguro	0.031	0.031	-0.028	0.093	-0.030	0.091
Seguro	0.012	0.019	-0.025	0.049	-0.025	0.048

Contraste de hipótesis

$$H_0 : p_A \leq p_B; \;\; H_1 : p_A > p_B$$

Expert	ph1.ini	ph1.post	b01
Inseguro	0.4967	0.8475	5.631222
Seguro	0.5080	0.7346	2.680720

Conclusiones y reflexiones