1 Teorema de De Finetti

Las urnas de Polya son un buen ejemplo de variables aleatorias intercambiables

2 Primeros ejemplos de medida de De Finetti

A pesar de que el teorema de representación de De Finetti nos garantizan la existencia de una distribución a priori no nos dice cual es el precisamente, podemos recolectar algunos ejemplos de ciertas bibliografías para hacernos una idea de como operar con estos conceptos. Si tenemos una muestra aleatoria independiente es simple ver que

$$f\left(x_1,\ldots,x_n\right)=\prod_{i=1}^{n} f\left(x_i\right) \implies f\left(x_{n+1},\ldots,x_{n+m}|x_1,\ldots,x_n\right)=\prod_{i=n+1}^{m} f\left(x_i\right)$$

, esto quiere decir que no hay aprendizaje desde la experiencia.

no hay aprendizaje desde la experiencia

Theorem 1. *Sea $\theta \sim f(\theta)$ y $Y_1,\ldots,Y_n$ condicionalmente independientes de θ entonces $Y_1,\ldots,Y_n$ es intercambiable.*

Proof. Basta con probar lo siguiente

$$\begin{aligned} f\left(y_1,\ldots,y_n\right) &= \int f\left(y_1,\ldots,y_n|\theta\right) dF\left(\theta\right) \\ &= \int \prod_{i=1}^{n} f\left(y_i|\theta\right) dF\left(\theta\right) \\ &= f\left(y_{\pi_1},\ldots,y_{\pi_n}\right) \end{aligned}$$

□

Básicamente el teorema de De Finetti es el recíproco de esta proposición. La prueba de De Finetti para variables aleatorias dicotómicas se explica a continuación. Consideremos el caso de una secuencia intercambiable de variables aleatorias $0-1$. Entonces

$$p\left(y_1,\ldots,y_n\right)=\int \prod_{i=1}^{n} p\left(y_i|\theta\right) dF\left(\theta\right)$$

donde

$$F\left(\theta\right)=\lim_{n\to\infty} P\left(s_n \leq \theta\right)$$

notemos que

$$p\left(y_1+\ldots+y_n=s_n\right)=\binom{n}{s_n} p\left(y_1,\ldots,y_n\right)=\binom{n}{s_n} p\left(y_{\pi_1},\ldots,y_{\pi_n}\right)$$

dado que vale para cualquier permutación, ahora bien, sea $N>n$ entonces

$$p\left(y_1+\ldots+y_n=s_n\right)=\sum p\left(y_1+\ldots+y_n=s_n|y_1+\ldots+y_N=s_N\right) p\left(y_1+\ldots+y_N=s_N\right)$$

2.1 Urnas de Polya

Las urnas de Polya son un buen ejemplo de intercambiabilidad: Tenemos b_0 bolitas negras y w_0 bolitas blancas en una urna. Extraigo una bolita aleatoriamente, si es blanca retorno la bolita a la urna y agrego una blanca, si es negra retorno la bolita y agrego una negra. B_n : El número de bolas negras dado n ensayos es.

$$B_n \sim Bin\left(n, \frac{b_0}{b_0 + w_0}\right)$$

la demostración biene como sigue

$$\begin{aligned} \text{Bolas negras iniciales } B_0 &= b_0 \\ \implies P\left(B_0 = b_0\right) &= 1 \\ P\left(B_1 = b_0 + 1\right) &= \frac{b_0}{b_0 + w_0} \end{aligned}$$

Para $n = 2$

$$\begin{aligned} P\left(B_2 = b_0\right) &= \frac{w_0}{w_0 + b_0} \cdot \frac{w_0 + 1}{w_0 + b_0 + 1} \\ P\left(B_2 = b_0 + 1\right) &= \frac{w_0}{w_0 + b_0} \frac{b_0}{w_0 + b_0 + 1} + \frac{b_0}{w_0 + b_0} \frac{w_0}{w_0 + b_0 + 1} \\ P\left(B_2 = b_0 + 2\right) &= \frac{b_0}{w_0 + b_0} \cdot \frac{b_0 + 1}{w_0 + b_0 + 1} \end{aligned}$$

debemos demostrar que la secuencia de variables aleatorias $X_1, \ldots, X_n$ de la secuencua de colores en cada una de las n extracciones en la urna es una secuencia de variables aleatorias intercambiables. Sea Y_n la proporción de bolas negras en la urna en dado n ensayos.

$$Y_n = \frac{B_n}{B_n + W_n} = \frac{B_n}{b_0 + w_0 + n}$$

con esto podemos decir que

$$B_{n+1} = \begin{cases} B_n & \text{con probabilidad } 1 - Y_n \\ B_n + 1 & \text{con probabilidad } Y_n \end{cases}$$

entonces

$$
\begin{aligned}
E\left(Y_{n+1}|X_1,\ldots,X_n\right) &= E\left(\frac{B_{n+1}}{b_0+w_0+n}|X_1,\ldots,X_n\right)\\
&= \frac{1}{b_0+w_0+n}E\left(B_{n+1}|X_1,\ldots,X_n\right)\\
&= \frac{1}{b_0+w_0+n}\left(B_n\left(1-Y_n\right)+\left(B_n+1\right)Y_n\right)\\
&= \frac{1}{b_0+w_0+n}\left(B_n-B_nY_n+B_nY_n+Y_n\right)\\
&= \frac{1}{b_0+w_0+n}\left(B_n+Y_n\right)\\
&= Y_n
\end{aligned}
$$

2.2 Blackwell and MacQueen 1973

Revisaremos la implicancias de este artículo en estadística bayesiana no paramétrica.

3 BDA 3

En esta sección se ilustrarán algunos conceptos útiles expuestos en el libro de Gelman y algunos ejercicios que son frecuentes en cátedras de inferencia bayesiana que se basan en el texto.

3.1 Introducció y motivación

3.2 Modelos uniparamétricos

Problem 2 (cap2.prob5)**.** La distribución a posterior es un compromiso entre la priori y los datos: Sea n el número de lanzamientos de una moneda, la cual tiene probabilidad de θ de ser cara. (a) Si la distribución a priori es una uniforme entre $0,1$, derive la distribución predictiva

$$p\left(y=k\right)=\int_0^1 p\left(y=k|\theta\right)d\theta$$

para cada $k = 1, \ldots, n$. Sol:

$$\begin{aligned}
p\left(y=k\right) &= \int_0^1 p\left(y=k,\theta\right) d\theta \\
&= \int_0^1 p\left(y=k|\theta\right) \underbrace{p\left(\theta\right)}_{1} d\theta \\
&= \int_0^1 p\left(y=k|\theta\right) d\theta \\
&= \int_0^1 \binom{n}{k} \theta^k \left(1-\theta\right)^{n-k} d\theta \\
&= \binom{n}{k} \int_0^1 \theta^k \left(1-\theta\right)^{n-k} d\theta \\
&= \binom{n}{k} \frac{\Gamma\left(k+1\right)\Gamma\left(n-k+1\right)}{\Gamma\left(n+2\right)} \\
&= \binom{n}{k} \frac{k!\left(n-k\right)!}{\left(n+1\right)!} \\
&= \frac{1}{n+1}
\end{aligned}$$

usamos el siguiente resultado

$$\begin{aligned}
&\int_0^1 \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)} \theta^{\alpha-1}\left(1-\theta\right)^{\beta-1} d\theta = 1 \\
&\quad \Longrightarrow \int_0^1 \theta^{\alpha-1}\left(1-\theta\right)^{\beta-1} d\theta = \frac{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta\right)} \\
&\quad\quad \Longrightarrow \int_0^1 \theta^k \left(1-\theta\right)^{n-k} d\theta = \frac{\Gamma\left(k+1\right)\Gamma\left(n-k+1\right)}{\Gamma\left(k+1+n-k+1\right)} = \frac{\Gamma\left(k+1\right)\Gamma\left(n-k+1\right)}{\Gamma\left(n+2\right)}
\end{aligned}$$

(b) Supona que asigna una distribución $\theta \sim \text{beta}\left(\alpha,\beta\right)$ y usted observa y caras en sus n lanzamientos. Muestre algebraicamente que la media posterior involucra a la media a priori y la media de los datos.

$$\begin{aligned}
p\left(\theta|y\right) &\propto p\left(y|\theta\right) p\left(\theta\right) \\
&= \binom{n}{y} \theta^y \left(1-\theta\right)^{n-y} \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)} \theta^{\alpha-1}\left(1-\theta\right)^{\beta-1} \\
&\propto \theta^{\alpha+y-1}\left(1-\theta\right)^{\beta+n-y-1} \\
\Longrightarrow\ & \theta|y \sim \text{beta}\left(\alpha+y, \beta+n-y\right)
\end{aligned}$$

luego, la media condicional

$$
\begin{aligned}
E(\theta|y) &= \frac{\alpha+y}{\alpha+y+\beta+n-y} \\
&= \frac{\alpha+y}{\alpha+\beta+n} \\
&= \frac{\alpha}{\alpha+\beta+n} + \frac{y}{\alpha+\beta+n} \\
&= \frac{\alpha+\beta}{\alpha+\beta}\frac{\alpha}{\alpha+\beta+n} + \frac{n}{n}\frac{y}{\alpha+\beta+n} \\
&= \frac{\alpha+\beta}{\alpha+\beta+n}\underbrace{\frac{\alpha}{\alpha+\beta}}_{E(\theta)} + \frac{n}{\alpha+\beta+n}\frac{y}{n}
\end{aligned}
$$

2.5c. Muestre que si la priori es uniforme entonces la media a posteriori, entonces la varianza a posterior es siempre menor a la varianza a priori. Sol: Sea $\theta \sim \text{beta}(1,1)$

$$
\begin{aligned}
\text{Var}(\theta|y) &= \frac{(\alpha+y)(\beta+n-y)}{(\alpha+y+\beta+n-y)^2(\alpha+y+\beta+n-y)} \\
&= \frac{(\alpha+y)(\beta+n-y)}{(\alpha+\beta+n)^2(\alpha+\beta+n+1)} \\
&= \frac{(1+y)(1+n-y)}{(2+n)^2(2+n+1)} = \frac{(1+y)(1+n-y)}{(2+n)^2(n+3)} \le \frac{1}{12} = \text{Var}(\theta)
\end{aligned}
$$

2.5d. De un ejemplo de una priori $\text{beta}(\alpha,\beta)$ con y y n donde va varianza posterior es menor a la varianza a priori. Sol.

$$
\frac{(\alpha+y)(\beta+n-y)}{(\alpha+\beta+n)^2(\alpha+\beta+n+1)} \le \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
$$

y seleccionamos valores que respeten la igualdad.

2.7a. Sobre prioris no informativas: Para la verosimilitud binomial $y \sim \text{bin}(n,\theta)$, muestre que $p(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ es la priori uniforme para el parámetro natural de la familia exponencial. Sol. Sabemos que la función de enlace canónica de distribución binomial es $\phi = \log\left(\frac{\theta}{1-\theta}\right) = h(\theta)$. sabiendo que

$$\begin{aligned} p(\theta) &= \underbrace{p(\phi)}_{\propto 1}\left|\frac{d\phi}{d\theta}\right| \\ &\propto \left|\frac{d}{d\theta}\log\left(\frac{\theta}{1-\theta}\right)\right| \\ &= \frac{1}{\theta} + \frac{1}{1-\theta} \\ &= \theta^{-1}(1-\theta)^{-1} \end{aligned}$$

.ver BDA (ecuación 2.19)

2.7b. Muestre que si $y = 0$ o $n = 0$ la distribución posterior es impropia. Sol: Si tenemos que

$$\begin{aligned} p(\theta|y) &\propto p(y|\theta)\, p(\theta) \\ &= \binom{n}{y}\theta^{y}(1-\theta)^{n-y}\,\theta^{-1}(1-\theta)^{-1} \\ &\propto \theta^{y-1}(1-\theta)^{n-y-1} = \begin{cases} \theta^{-1}(1-\theta)^{n-1} & y = 0 \\ \theta^{-1}(1-\theta)^{-1} & n = 0 \\ \theta^{y-1}(1-\theta)^{n-y-1} & \text{otro caso} \end{cases} \end{aligned}$$

en los dos primeros casos no integra 1

2.8a. Distribución normal con media desconocida: Sea una muestra aleatoria de n estudiantes extraidos de una población lo suficientemente grande para luego medir su peso. La media del peso viene dada por $\overline{y} = 150$. Asuma que pero en la población se distribuye normal con media desconocida θ y desviación estandar de 20. Suponga una distribución a priori para θ normal con media 180 y desviación estandar de 40. Encuentre la distribución a posterior de θ. Sol: Solo con el objetivo de ejercitar todos los pasos deduciremos la distribución a posterior completamente.

$$\begin{aligned} y|\theta &\sim N\left(\theta, \sigma^2\right) \\ \theta &\sim N\left(\mu_0, \tau_0^2\right) \end{aligned}$$

entonces tenemos que

$$\begin{aligned}
p\left(\theta|y\right) &\propto p\left(y|\theta\right)p\left(\theta\right)\\
&=\prod_{i=1}^{n}\left[\left(2\pi\sigma^2\right)^{-1/2}\exp\left(-\frac{1}{2\sigma^2}\left(y_i-\theta\right)^2\right)\right]\left(2\pi\tau_0^2\right)^{-1/2}\exp\left(-\frac{1}{2\tau_0^2}\left(\theta-\mu_0\right)^2\right)\\
&\propto \exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-\theta\right)^2\right)\exp\left(\frac{1}{2\tau_0^2}\left(\theta-\mu_0\right)^2\right)\\
&=\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i-\theta\right)^2-\frac{1}{2\tau_0^2}\left(\theta-\mu_0\right)^2\right)\\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i^2-2y_i\theta+\theta^2\right)-\frac{1}{2\tau_0^2}\left(\theta^2-2\theta\mu_0+\mu^2\right)\right)\\
&\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\theta^2-2y_i\theta\right)-\frac{1}{2\tau_0^2}\left(\theta^2-2\theta\mu_0\right)\right)
\end{aligned}$$

entonces

$$\theta|y\sim N\left(\frac{\frac{1}{\sigma^2}\mu_0+\frac{n}{\tau_0^2}\overline{y}}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}},\frac{1}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}}\right)$$

para nuestro caso tenemos que

$$\theta|y\sim N\left(\frac{\frac{1}{\sigma^2}\mu_0+\frac{n}{\tau_0^2}\overline{y}}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}},\frac{1}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}}\right)$$

en nuestro cado

$$\theta|y\sim N\left(\frac{\frac{1}{20^2}180+\frac{n}{40^2}150}{\frac{1}{20^2}+\frac{n}{40^2}},\frac{1}{\frac{1}{20^2}+\frac{n}{40^2}}\right)$$

2.8b. La forma de encontrar la distribución predictiva en media esperanza y varianza iterada.

$$\tilde{y}|y\sim N\left(\widetilde{\mu},\widetilde{\sigma}^2\right)$$

donde

$$\begin{aligned}
\widetilde{\mu}&=E\left(\tilde{y}|y\right)\\
&=E\left(E\left(\tilde{y}|\theta,y\right)|y\right)\\
&=E\left(E\left(\tilde{y}|\theta\right)|y\right)\\
&=E\left(\theta|y\right)\\
&=\frac{\frac{1}{\sigma^2}\mu_0+\frac{n}{\tau_0^2}\overline{y}}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}}
\end{aligned}$$

y la varianza

$$\begin{aligned}\widetilde{\sigma}^2 &= var\left(\tilde{y}|y\right)\\ &= var\left(E\left(\tilde{y}|\theta\right)|y\right) + E\left(var\left(\tilde{y}|\theta\right)|y\right)\\ &= var\left(\theta|y\right) + E\left(\sigma^2|y\right)\\ &= \frac{1}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}} + \sigma^2\end{aligned}$$

por lo tanto

$$\tilde{y}|y \sim N\left(\frac{\frac{1}{20^2}180 + \frac{n}{40^2}150}{\frac{1}{20^2}+\frac{n}{40^2}}, \frac{1}{\frac{1}{\sigma^2}+\frac{n}{\tau_0^2}} + \sigma^2\right)$$

4.2.a normalidad asintótica de la posterior: sea $y_1 \ldots, y_5$ muestras independientes de una distribución cauchy con parametro de posición desconocido θ y escala conicda 1. $p\left(y_i|\theta\right) \propto \frac{1}{1+(y_i-\theta)^2}$. Asuma que la distribución a priori para

$$\theta \sim \text{uniforme}\left(0,1\right)$$

determine la derivada y la segunda derivada del logaritmo de la verosimilitud.
sol

$$l\left(\theta|y\right) = \prod_{i=1}^{n} \frac{1}{1+\left(y_i-\theta\right)^2}$$

4 Ayudantía

Ejercicios de las ayudantías del segundo semestre del 2020

4.1 Ayudantía 1

4.2 Ayudantía 2

4.3 Ayudantía 3

4.4 Ayudantía 4

4.5 Ayudantía 5

4.5.1a. x_i permutables $p\left(\mu,\sigma^2\right)\propto\left(\sigma^2\right)^{-3/2}$. Buscamos la distribución a posterior

$$\begin{aligned} p\left(\mu,\sigma^2|\mathbf{x}\right) &\propto p\left(\mu,\sigma^2\right)p\left(\mathbf{x}|\mu,\sigma^2\right)\\ &=\left(\sigma^2\right)^{-3/2}\sigma^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum\left(x_i-\mu\right)^2\right)\\ &=\left(\sigma^2\right)^{-3/2}\sigma^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum\left(x_i-\overline{x}\right)^2+n\left(\overline{x}-\mu\right)^2\right)\\ &=\left(\sigma^2\right)^{-3/2}\sigma^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\left(\sum\left(x_i-\overline{x}\right)^2+n\left(\overline{x}-\mu\right)^2\right)\right)\\ &=\left(\sigma^2\right)^{-3/2-n/2}\exp\left(-\frac{1}{2\sigma^2}\left(\left(n-1\right)s^2+n\left(\overline{x}-\mu\right)^2\right)\right)\end{aligned}$$

kerner de una normal $\propto\exp\left(-\frac{1}{2\sigma^2}\left(n\left(\overline{x}-\mu\right)^2\right)\right)$

$$\implies \mu|\mathbf{x},\sigma^2\sim N\left(\overline{x},\sigma^2/n\right)$$

4.5.1b. posteriori marginal

$$\begin{aligned} p\left(\mu,\sigma^2|\mathbf{x}\right) &\propto \left(\sigma^2\right)^{-3/2-n/2}\exp\left(-\frac{1}{2\sigma^2}\left(\left(n-1\right)s^2+n\left(\overline{x}-\mu\right)^2\right)\right)\\ &=\left(\sigma^2\right)^{-3/2-n/2}\exp\left(-\frac{1}{2\sigma^2}\left(n-1\right)s^2\right)\underbrace{\exp\left(-\frac{1}{2\sigma^2}n\left(\overline{x}-\mu\right)^2\right)\left(\frac{\sigma}{n}\right)^{-1/2}}_{\text{kernel de una normal}}\left(\frac{\sigma}{n}\right)^{1/2}\end{aligned}$$

integrando

$$\begin{aligned}\implies p\left(\sigma^2|\mathbf{x}\right) &\propto \left(\sigma^2\right)^{-(n/2+1)}\exp\left(-\frac{1}{2\sigma^2}\left(n-1\right)s^2\right)\\ \sigma^2|\mathbf{x} &\sim \chi^2-\mathbf{inv}\left(n,s^2\right)\end{aligned}$$

4.5.1c. calcular la distribución a posteriori marginal de μ

$$
\begin{aligned}
p\left(\mu|\mathbf{x}\right) &= \int_0^\infty p\left(\mu,\sigma^2|\mathbf{x}\right) d\sigma^2 \\
&\propto \int_0^\infty \left(\sigma^2\right)^{-3/2} \sigma^{-n/2} \exp\left(-\frac{1}{2\sigma^2}A\right) d\sigma^2 \\
A &= \sum \left(x_i-\overline{x}\right)^2 + n\left(\overline{x}-\mu\right)^2
\end{aligned}
$$

con el siguiente cambio de variable $z = A/\left(2\sigma^2\right)$, $\sigma^2 = A/2z$ y $d\sigma^2 = -\frac{A}{2z^2}dz$. Luego

$$
\begin{aligned}
p\left(\mu|\mathbf{x}\right) &\propto \int_0^\infty \left(\frac{A}{2z}\right)^{-3/2-n/2} \exp\left(-z\right)\left(-\frac{A}{2z^2}\right) dz \\
&\propto \left(\frac{A}{2z}\right)^{-1/2-n/2}
\end{aligned}
$$

4.5.1d. Utilice los resultados anteriores para encontrar la distribución predictiva a posteriori usando el método de la composición.

A) para $s = 1\ldots,S$

A) genera $\mu^{(s)} \sim p\left(\mu|\mathbf{y}\right)$

A) genera $\sigma^{2(s)} \sim p\left(\sigma^2|\mathbf{y}\right)$

A) $y^s \sim p\left(x|\mu^{(s)},\sigma^{2(s)}\right)$

4.5.1e Repita a) y c) con la priori informativa: $p\left(\mu,\sigma^2\right) = p\left(\mu|\sigma^2\right)p\left(\sigma^2\right)$. Luego

$$
\begin{aligned}
p\left(\mu,\sigma^2|\mathbf{x}\right) &\propto p\left(\mu,\sigma^2\right)p\left(\mathbf{x}|\mu,\sigma^2\right) \\
&= p\left(\mu|\sigma^2\right)p\left(\sigma^2\right)p\left(\mathbf{x}|\mu,\sigma^2\right) \\
&=
\end{aligned}
$$

4.5.2a. Encontre la distribución a posterior de β, σ^2. Conocido que la distribución a priori de $p\left(\beta, \sigma^2\right) \propto \sigma^{-2}$

$$
\begin{aligned}
y|\beta, \sigma^2, X &\sim N\left(X\beta, \sigma^2 I\right) \\
p\left(y|\beta, \sigma^2, X\right) &\propto \det\left(\sigma^2 I\right)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right) \\
&= \left(\sigma^2\right)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right) \\
&= \\
\implies p\left(\beta, \sigma^2|y, X\right) &\propto \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right)\sigma^{-2} \\
&= \left(\sigma^2\right)^{-n/2-2} \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right)
\end{aligned}
$$

4.5.2b. Demuestre que la distribución a posterior de β dado σ^2 corresponde a una distribución normal $\hat{\beta}, \sigma^2\left(X^t X\right)^{-1}$**.** Sol: Sabemos que

$$
\begin{aligned}
p\left(\beta, \sigma^2|y\right) &= p\left(\sigma^2|\beta, y\right) p\left(\beta|y\right) && \implies p\left(\beta|y\right) = \frac{p\left(\beta, \sigma^2|y\right)}{p\left(\sigma^2|\beta, y\right)} = \frac{p\left(y|\beta, \sigma^2\right) p\left(\beta, \sigma^2\right)}{p\left(\sigma^2|\beta, y\right)} \\
p\left(\beta, \sigma^2|y\right) &= p\left(\beta|\sigma^2, y\right) p\left(\sigma^2|y\right) && \implies p\left(\sigma^2|y\right) = \frac{p\left(y|\beta, \sigma^2\right) p\left(\beta, \sigma^2\right)}{p\left(\beta|\sigma^2, y\right)}
\end{aligned}
$$

El camino quizás es más simple si hacemos la siguiente operatoria:

$$
\begin{aligned}
p\left(\beta|\sigma^2, y\right) &= \frac{p\left(\beta, \sigma^2|, y\right)}{p\left(\sigma^2|y\right)} \\
&\propto p\left(\beta, \sigma^2|, y\right) \\
&= \left(\sigma^2\right)^{-n/2-2} \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\left(y - X\beta\right)^t\left(y - X\beta\right)\right) \\
&= \exp\left(-\frac{1}{2\sigma^2}\left(y - X\hat{\beta}\right)^t\left(y - X\hat{\beta}\right) - \frac{1}{2\sigma^2}\left(\beta - \hat{\beta}\right)^t X^t X\left(\beta - \hat{\beta}\right)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\left(\beta - \hat{\beta}\right)^t X^t X\left(\beta - \hat{\beta}\right)\right) \\
\implies \beta|\sigma^2, y &\sim N\left(\hat{\beta}, \sigma^2\left(X^t X\right)^{-1}\right)
\end{aligned}
$$

4.5.2c. Encuentre la distribución a posterior de σ^2 , (ver ejercicio 14.4 BDA)

$$\begin{aligned}
p\left(\sigma^2|y\right) &= \frac{p\left(\beta,\sigma^2|,y\right)}{p\left(\beta|\sigma^2,y\right)} \\
&\propto \frac{\left(\sigma^2\right)^{-n/2-2}\exp\left(-\frac{1}{2\sigma^2}\left(y-X\beta\right)^t\left(y-X\beta\right)\right)}{\det\left(\sigma^2\left(X^tX\right)^{-1}\right)^{-1/2}\exp\left(-\frac{1}{2\sigma^2}\left(\beta-\hat{\beta}\right)^tX^tX\left(\beta-\hat{\beta}\right)\right)} \\
&= \frac{\left(\sigma^2\right)^{-n/2-2}\exp\left(-\frac{1}{2\sigma^2}\left(y-X\hat{\beta}\right)^t\left(y-X\hat{\beta}\right)-\frac{1}{2\sigma^2}\left(\beta-\hat{\beta}\right)^tX^tX\left(\beta-\hat{\beta}\right)\right)}{\det\left(\sigma^2\left(X^tX\right)^{-1}\right)^{-1/2}\exp\left(-\frac{1}{2\sigma^2}\left(\beta-\hat{\beta}\right)^tX^tX\left(\beta-\hat{\beta}\right)\right)} \\
&= \frac{\left(\sigma^2\right)^{-n/2-2}\exp\left(-\frac{1}{2\sigma^2}\left(y-X\hat{\beta}\right)^t\left(y-X\hat{\beta}\right)\right)}{\det\left(\sigma^2\left(X^tX\right)^{-1}\right)^{-1/2}\exp\left(0\right)} \\
&= \frac{\left(\sigma^2\right)^{-n/2-2}\exp\left(-\frac{1}{2\sigma^2}\left(y-X\hat{\beta}\right)^t\left(y-X\hat{\beta}\right)\right)}{\left(\sigma^{2k}\right)^{-1/2}\det\left(\left(X^tX\right)^{-1}\right)^{-1/2}\exp\left(0\right)} \\
&\propto \left(\sigma^2\right)^{-n/2-2+k/2}\exp\left(-\frac{1}{2\sigma^2}\left(y-X\hat{\beta}\right)^t\left(y-X\hat{\beta}\right)\right)
\end{aligned}$$

$$\implies \sigma^2|y \sim Inv-\chi^2\left(n-k,s^2\right)$$

4.5.2d Si $\widetilde{y}$ es una nueva respuesta, existen a y b tales que

5 Ejemplos de cátedra

demuestre que que modelo probit con $\pi\left(x\right)=\Phi\left(\beta_0+\beta_1x|\mu,\sigma^2\right)$ es no identificado para μ y σ^2. La demostración se basa en que

$$\begin{aligned}
& Y_i|x_i \sim \text{ber}\left(\pi\left(x_i\right)\right) \\
\implies & \pi\left(x\right)=\Phi\left(\beta_0+\beta_1x|\mu,\sigma^2\right) \\
& \pi\left(x\right)=\Phi_{0,1}\left(\frac{\beta_0+\beta_1x-\mu}{\sigma}\right) \\
& \pi\left(x\right)=\Phi_{0,1}\left(\beta_0^*+\beta_1^*x\right)
\end{aligned}$$

entonces existen parámetros observacionalmente equivalentes.

Intervalo de Confianza El parámetro está cubierto por el intervalo de confianza con una probabilidad $1-\alpha$.

Identificabilidad bayesiana Si tengo un modelo clasicamente no identificado, en bayesiana podemos