INFERENCIA BAYESIANA

contenidos:
- conceptos basicos
- aplicación a los modelos
  - → modelo Beta-Binomial
  - → modelo Poisson
  - → modelo Normal
- modelos multiparametricos

1) CONCEPTOS BASICOS

Thm: Bayes theorem

TEOREMA DE BAYES

- Para eventos: $(\Omega, \mathcal{A}, \mathbb{P})$ espacio de probabilidad,
  $\{A_i\}_{i=1}^{n}$ succesion de eventos con prob. no nulos
  B evento de prob. no nulo

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}(A_k)\mathbb{P}(B \mid A_k)}{\sum_{i=1}^{n} \mathbb{P}(A_i)\mathbb{P}(B \mid A_i)} \quad \forall k = 1, \dots, n$$

- Para variables aleatorios discretas

X v.a. $\underset{\text{discr.}}{\rightsquigarrow}$ $\operatorname{Im}(X) = \{x_1, \dots, x_m\} \longrightarrow$ partición $\{A_j\}_{j=1}^{m}$

Y v.a. $\rightsquigarrow$ $\operatorname{Im}(Y) = \{y_1, \dots, y_n\} \longrightarrow$ partición $\{B_i\}_{i=1}^{n}$

$$\mathbb{P}(Y = y_i \mid X = x_j) = \frac{\mathbb{P}(Y = y_i)\mathbb{P}(X = x_j \mid Y = y_i)}{\sum_{k=1}^{n} \mathbb{P}(Y = y_k)\mathbb{P}(X = x_j \mid Y = y_k)} \quad \begin{array}{l}\forall i = 1, \dots, n \\ \forall j = 1, \dots, m\end{array}$$

$$\left(P(A_i) = P(w \in \Omega : X(w) = x_i) = P(X = x_i) \right)$$

- Para variables aleatorias continuas

(X, Y) vect. aleatorio bidim. abs. cont. con

→ $f(x,y)$ función de densidad

→ h densidad marginal de X

→ g densidad marginal de Y

$$\Rightarrow f(y|x) = \frac{f(x,y)}{h(x)} = \frac{g(y)\, f(x|y)}{h(x)}$$

$$= \frac{g(y)\, f(x|y)}{\int f(x,y)\, dy} = \frac{g(y)\, f(x|y)}{\int g(y)\, f(x|y)\, dy}$$

ESTIMACIÓN

$$\theta \longrightarrow \underline{x} = (x_1, \dots, x_n) \longrightarrow \theta | \underline{x}$$

APRIORI DATOS APOSTERIORI

def: $\mathcal{L} : \Theta \times \Theta \longrightarrow \mathbb{R}^+$ FUNCIÓN DE PERDIDA

~ función que cuantifica el error de estimación de $\hat{\theta}$ (ESTIMADOR) con respecto a θ

ej:
- $\mathcal{L}(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$ CUADRATICA
- $\mathcal{L}(\hat{\theta}, \theta) = |\hat{\theta} - \theta|$ norma L_1

Para demostrar los valores optimo por las funciones de perdida previos necesitamos un resultado preliminar

lemma: X v.a. , $X \geq 0$, $\mathbb{E}[X] < +\infty$

$$\Rightarrow \quad \mathbb{E}[X] = \int_0^{+\infty} \left(1 - F_X(t)\right) dt$$

dim: $$\mathbb{E}[X] = \lim_{m \to +\infty} \int_0^{+\infty} x \, dF(x) \underset{IPP}{=} \lim_{m \to +\infty} \left[x\left[F(x)\right]_0^m - \int_0^m F(x)\,dx \right]$$

$$= \lim_{m \to +\infty} \left[m F(m) - \int_0^m F(x)\,dx \right]$$

$$= \lim_{m \to +\infty} \left[m\left(F(m) - 1\right) + \int_0^m \left(1 - F(t)\right) dt \right]$$

Ahora demostramos que $\lim_{m \to +\infty} \left[m\left(F(m) - 1\right) \right] = 0$

$$\mathbb{E}[X] = \lim_{m \to +\infty} \left(\int_0^m x \, dF(x) + \int_m^{+\infty} x \, dF(x) \right)$$

$$\geq \lim_{m \to +\infty} \left(\int_0^m x \, dF(x) + m \, \mathbb{P}(X \geq m) \right)$$

$$= \mathbb{E}[X] + \underbrace{\lim_{m \to +\infty} m \, \mathbb{P}(X \geq m)}_{=0}$$

N.B $\quad m\,\mathbb{P}(X \geq m) = m\left(1 - F(m)\right)$

$$\Rightarrow \quad \mathbb{E}[X] = \int_0^{+\infty} \left(1 - F(t)\right) dt \qquad \square$$

prop: ESTIMADORES DE BAYES

función de pérdida $\longrightarrow$ estimador

1. cuadrática $\longrightarrow$ media aposteriori
2. norma L_1 $\longrightarrow$ mediana aposteriori

dem:

1) definimos $g(\delta) = \mathbb{E}[(\theta - \delta)^2 \mid \underline{x}]$

estimador puntual de δ (dependiente de los datos)

$$\rightarrow g'(\delta) = 2\left(\mathbb{E}[\theta \mid \underline{x}] - \delta\right) = 0 \quad (\text{min})$$

$$\Rightarrow \delta = \mathbb{E}[\theta \mid \underline{x}]$$

$$\left[\frac{\partial}{\partial \delta} \mathbb{E}[\theta^2 - 2\theta\delta + \delta^2 \mid \underline{x}] = 2\Big(\underbrace{\mathbb{E}[\delta \mid \underline{x}]}_{\delta} - \mathbb{E}[\theta \mid \underline{x}]\Big) \right]$$

2) $$\mathbb{E}[\,|\theta - \delta| \mid \underline{x}] = \int_{\delta}^{+\infty} (1 - \mathbb{P}(\theta \leq t \mid \underline{x}))\,dt - \int_{-\infty}^{\delta} \mathbb{P}(\theta \leq t \mid \underline{x})\,dt$$

$$\hookrightarrow \frac{\partial}{\partial \delta} \longrightarrow 1 - 2\mathbb{P}(\theta \leq \delta \mid \underline{x}) = 0$$

$$\Rightarrow \delta \text{ t.q. } \mathbb{P}(\theta \leq \delta \mid \underline{x}) = \frac{1}{2}$$

es decir $\delta = \text{Median}(\theta \mid \underline{x})$ $\square$

INTERVALOS DE CREDIB.

def: $\gamma \in (0,1)$, nivel de CREDIBILIDAD

(t_l, t_u) t.q. $\displaystyle\int_{t_l}^{t_u} f(\theta \mid \underline{x})\, d\theta = \gamma$

con $\underline{x} = (x_1, \ldots, x_n)$ datos

INTERVALO DE CREDIBILIDAD (nivel γ)

metodos de calculo:

(1) para distr. unimodales: HPDI (high posterior density interval)

→ calculamos un intervalo de prob. que contenga a la moda y t.q. cada punto de este intervalo tenga valor de prob. mayor que cualquier otro punto fuera

(2) para contener la mediana: se calcula repartiendo la la misma probabilidad en las colas

(METODO DE COLAS IGUAL. POND.)

(3) si la media existe, intervalo t.q. media es punto central

---

ELECCIÓN DE APRIORI

metodos de elección de la apriori:

→ distribuciones conjugadas

→ distribuciones no informativas (impropias)

→ apriori de Jeffrey

def: DISTRUBUCIONES CONJUGADAS

$L(\theta) = f(\underline{x} \mid \theta)$ verosimilitud de θ

$\mathcal{C}$, clase de distr., se define CONJUGADA con respecto a $L(\theta)$, si para una apriori $\pi(\theta) \in \mathcal{C}$ la distr. aposteriori $\pi(\theta \mid \underline{x}) \in \mathcal{C}$ $\forall \underline{x}$

Es decir, $\boxed{\text{si } \pi(\theta) \in \mathcal{C} \Rightarrow \pi(\theta \mid \underline{x}) \in \mathcal{C} \quad \forall \underline{x}}$

def: DISTRIBUCIONES IMPROPIAS

Una función t.q. no cumple la propriedad $\int_\Omega f(x)\,dx = 1$ se define DISTR. IMPROPIA

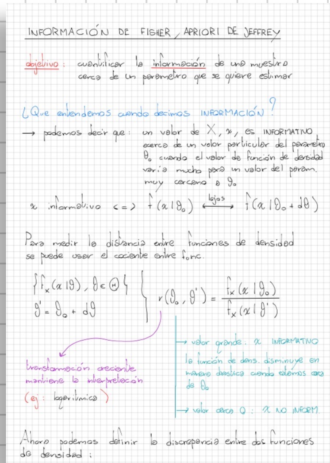
def: DISTRIBUCIÓN DE JEFFREY (para el caso unidimensional)

$$\pi(\theta) \propto I^{1/2}(\theta)$$

acuerdo: $$I(\theta) = \mathbb{E}_\theta\left[\left(\frac{\partial \log f(X \mid \theta)}{\partial \theta}\right)^2\right]$$

INFORMACIÓN DE FISHER

excursus: porque la apriori de Jeffrey?



INFORMACIÓN DE FISHER, APRIORI DE JEFFREY

objetivo: cuantificar la información de una muestra cerca de un parametro que se quiere estimar

¿Que entendemos cuando decimos INFORMACIÓN?

→ podemos decir que: un valor de X, x, es INFORMATIVO acerca de un valor particular del parametro θ_0 cuando el valor de función de densidad varía mucho para un valor del param. muy cercano a θ_0

x informativo $\iff f(x \mid \theta_0) \xleftrightarrow{\text{lejos}} f(x \mid \theta_0 + d\theta)$

Para medir la distancia entre funciones de densidad se puede usar el cociente entre func.

$$\left.\begin{array}{l} \{f_X(x \mid \theta),\ \theta \in \Theta\} \\ \theta' = \theta_0 + d\theta \end{array}\right\} \quad r(\theta_0, \theta') = \frac{f_X(x \mid \theta_0)}{f_X(x \mid \theta')}$$

transformación creciente mantiene la interpretación (ej: logarítmica)

→ valor grande: x INFORMATIVO

la función de dens. disminuye en manera drastica cuando estamos cerca de θ_0

→ valor cerca 0: x NO INFORM.

Ahora podemos definir la discrepancia entre dos funciones de densidad:

objetivo: Dado una muestra $\underline{x} = (x_1, \dots, x_n)$ predecir

$$y \sim f(\cdot \mid \theta)$$

DISTRIBUCIÓN PREDICTIVA

Hay dos maneras de calcular la distr. de y

(1) aplicando el thm de Bayes

$$\pi(\theta \mid y, \underline{x}) = \frac{\pi(\theta \mid \underline{x}) f(y \mid \theta, \underline{x})}{f(y \mid \underline{x})}$$

$$f(y \mid \underline{x}) = \frac{\pi(\theta \mid \underline{x}) f(y \mid \theta, \underline{x})}{\pi(\theta \mid y, \underline{x})}$$

(2) ponderación con la aposteriori

$$f(y \mid \underline{x}) = \int_{\Theta} f(y \mid \theta, \underline{x}) \pi(\theta \mid \underline{x}) \, d\theta$$

$$\begin{cases} H_0: \theta \in \Theta_0 \\ H_1: \theta \in \Theta_1 \end{cases}$$

a partir de una muestra $\underline{x}$

FACTOR DE BAYES

def: $$FB_{01} = \frac{\mathbb{P}(\underline{x} \mid H_0)}{\mathbb{P}(\underline{x} \mid H_1)}$$ FACTOR DE BAYES interpretación: FACTOR DE BAYES

si $FB_{01} \approx 0 \Rightarrow P(\underline{x} \mid H_0) << P(\underline{x} \mid H_1)$

evidencia para decir que los datos son generatos por H_1

si FB_{01} grande $\Rightarrow P(\underline{x} \mid H_0) >> P(\underline{x} \mid H_1)$

evidencia para decir que los datos son generatos por H_0

Ahora aplicando Bayes:

$$P(\underline{x} \mid H_0) = \frac{P(\underline{x})\, P(H_0 \mid \underline{x})}{P(H_0)}$$

$$P(\underline{x} \mid H_1) = \frac{P(\underline{x})\, P(H_1 \mid \underline{x})}{P(H_1)}$$

$$\hookrightarrow FB_{01} = \frac{P(H_1 \mid \underline{x})}{P(H_0 \mid \underline{x})} \cdot \frac{P(H_0)}{P(H_1)}$$

2) APLICACIÓN A LOS MODELOS

MODELO BETA-BINOMIAL

datos

$\underline{x} = (x_1, \dots, x_n)$ donde $x_i \in \{0,1\}$ $\forall i$

- función de verosimilitud: $(X \sim Be(p))$

$$L(p \mid \underline{x}) = f(\underline{x} \mid p) = \prod_{i=1}^{n} f(x_i \mid p)$$
$$= \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$
$$= p^{\Sigma x_i} (1-p)^{n - \Sigma x_i}$$

- apriori $\longrightarrow$ aposteriori

→ NO INFORMATIVA COMPLETA: $\pi(p) = I_{[0,1]}(p)$

$(p \sim Beta(1,1) = U_{[0,1]})$

$$\pi(p \mid \underline{x}) \propto L(p \mid \underline{x})\, \pi(p) = p^{\Sigma x_i} (1-p)^{n - \Sigma x_i}$$

$$\hookrightarrow \boxed{p \mid \underline{x} \sim Beta\left(\Sigma x_i + 1,\ n - \Sigma x_i + 1\right)}$$

→ APRIORI DE JEFFREY: $\pi(p) = p^{-1/2} (1-p)^{-1/2}$

Calculamos la inform. de Fisher

$$f(x|p) = p^{x}(1-p)^{1-x} \longrightarrow \ln f(x|p) = x\ln(p) + (1-x)\ln(1-p)$$

$$\frac{\partial \ln f(x|p)}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p} = \frac{x(1-\cancel{p}) - p(1-\cancel{x})}{p(1-p)} = \frac{x-p}{p(1-p)}$$

$$\hookrightarrow I(p) = \mathbb{E}\left[\left(\frac{\partial \ln f}{\partial p}\right)^2\right] = \frac{1}{p^2(1-p)^2}\,\mathbb{E}\left[(x-p)^2\right]$$

$$= \frac{1}{p(1-p)}$$

$$\Rightarrow \pi(p) = p^{-1/2}(1-p)^{-1/2}$$

$$\Rightarrow p \sim \text{Beta}\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$$\Rightarrow \pi(p|\underline{x}) = p^{\sum x_i - \frac{1}{2}}(1-p)^{n-\sum x_i - \frac{1}{2}}$$

$$\Rightarrow \boxed{p|\underline{x} \sim \text{Beta}\left(\sum x_i + \tfrac{1}{2},\ n - \sum x_1 + \tfrac{1}{2}\right)}$$

→ APRIORI POR EL METODO DE LAS CONJUGADAS

$$p \sim \text{Beta}(a, b)$$

$$\Rightarrow \boxed{p|\underline{x} \sim \text{Beta}\left(\sum x_i + a + 1,\ n - \sum x + b + 1\right)}$$

$$p) = p^{\sum x_i}(1-p)$$

- estimadores

Suponemos una aposteriori: $p \mid \underline{x} \sim \text{Beta}(a(\underline{x}), b(\underline{x}))$

Función de perdida cuadr.: $\mathbb{E}[p \mid \underline{x}] = \dfrac{a(\underline{x})}{a(\underline{x}) + b(\underline{x})}$

Función de perd. L_1: $\text{Med}(p \mid \underline{X}) = \dots$

↳ no calculo expl.

- distribución predictiva: $Y \sim \text{Bern}(p^*)$

$$p^* = \mathbb{P}(Y = 1 \mid \underline{x}) = \int_0^1 P(Y=1 \mid \underline{x}, p)\, \pi(p \mid \underline{x})\, dp$$

generica apriori Beta(a,b)

$$= \int_0^1 p \cdot \frac{p^{\sum x_i + a + 1} (1-p)^{n - \sum x_i + b + 1}}{B(\sum x_i + a, n - \sum x_i + b)}\, dp$$

$$= \frac{B(\sum x_i + a + 1, n - \sum x_i + b)}{B(\sum x_i + a, n - \sum x_i + b)}$$

MODELO POISSON

$\underline{x} = (x_1, \dots, x_n)$ con $x_i \in \mathbb{N} \ \forall i$

- función de verosimilitud

$$L(\lambda | \underline{x}) = \prod_{i=1}^{n} f(x_i | \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \propto \lambda^{\Sigma x_i} e^{-n\lambda}$$

- apriori $\longrightarrow$ aposterior:

→ APRIORI DE JEFFREY

Calculamos la información de Fisher:

$$\ln f(x|\lambda) = \ln\left(e^{-\lambda} \frac{\lambda^x}{x!}\right) = -\lambda + x \ln \lambda - \ln(x!)$$

$$\frac{\partial \ln f(x|\lambda)}{\partial \lambda} = \frac{x}{\lambda} - 1$$

$$\Rightarrow \quad I(\lambda) = \mathbb{E}\left[\left(\frac{x-\lambda}{\lambda}\right)^2\right] = \frac{1}{\lambda^2} \cdot \lambda = 1/\lambda$$

$$\Rightarrow \pi(\lambda) = 1/\sqrt{\lambda} = \lambda^{-1/2}$$

$$\Rightarrow \pi(\lambda | \underline{x}) \propto \lambda^{\Sigma x_i - 1/2} e^{-n\lambda}$$

es decir $\boxed{\lambda | \underline{x} \sim \text{Gamma}\left(\Sigma x_i + \frac{1}{2}, n\right)}$

→ APRIORI CON EL METODO DE LAS CONJUGADAS

$$\lambda \sim \text{Gamma}(\alpha, \beta)$$

$$\Rightarrow \pi(\lambda|\underline{x}) \propto \lambda^{\sum x_i} e^{-n\lambda} \cdot \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$\propto \lambda^{\sum x_i + \alpha - 1} \cdot e^{-(n+\beta)\lambda}$$

$$\Rightarrow \lambda|\underline{x} \sim \text{Gamma}\left(\sum x_i + \alpha, n+\beta\right)$$

→ APRIORI NO INFORMATIVA COMPLETA

$$\pi(\lambda) = I_{[0,+\infty)}(\lambda)$$

$$\Rightarrow \pi(\lambda|\underline{x}) = \lambda^{\sum x_i} e^{-n\lambda}$$

$$\Rightarrow \lambda|\underline{x} \sim \text{Gamma}\left(\sum x_i + 1, n\right)$$

- estimadores

Asumendo como aposteriori: $\lambda|\underline{x} \sim \text{Gamma}\left(\alpha(\underline{x}), \beta(\underline{x})\right)$

- función de perdida cuadr. $E[\lambda|\underline{x}] = \dfrac{\alpha(\underline{x})}{\beta(\underline{x})}$
- función de perdida L_1: mediana ⇝ calculada con R

- distribución predictiva

$\alpha = \alpha(\underline{x}),\ \beta = \beta(\underline{x})$

$$Y \sim \text{Pois}(\lambda), \quad \lambda|\underline{x} \sim \text{Gamma}(\alpha, \beta)$$

$$\Rightarrow P(Y=k|\underline{x}) = \int_0^{+\infty} P(Y=k|\lambda,\underline{x})\,\pi(\lambda|\underline{x})\,d\lambda$$

$$= \int_0^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot \beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda} \cdot \frac{1}{\Gamma(\alpha)}\,d\lambda$$

$$= \frac{1}{k!} \beta^{\alpha} \cdot \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} e^{-(\beta+1)\lambda} \lambda^{k+\alpha-1} \, d\lambda$$

$$= \frac{1}{k!} \cdot \frac{\beta^{\alpha}}{(\beta+1)^{\alpha+k}} \cdot \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \quad \text{si } \alpha \in \mathbb{N}$$

$$\hookrightarrow = \frac{(\alpha+k-1)!}{(\alpha-1)!} = (\alpha+k-1)(\alpha+k-2)\cdots\alpha$$

$$= \underbrace{\frac{(\alpha+k-1)!}{k!(\alpha-1)!}}_{\binom{\alpha+k-1}{k}} \cdot \left(\frac{\beta}{\beta+1}\right)^{\alpha} \left(\frac{1}{\beta+1}\right)^{k}$$

es decir $Y \mid \underline{x} \sim NB\left(\frac{\beta}{\beta+1}, n\right)$

DISTRIBUCIÓN BINOMIAL NEGATIVA

$X \sim NB(p, n)$ ~ numero de fracasos antes de n-sucesos

$$\mathbb{P}(X = k \mid p, n) = \binom{k+n}{k-1} p^{k}$$

$$= \binom{-n}{k} p^{n} (p-1)^{k}$$

$$\mathbb{E}[X \mid p, n] = \frac{n}{p}, \quad Var(X \mid p, n) = n\frac{(1-p)}{p^2}$$

MODELO NORMAL (varianza conocida)

$\underline{x} = (x_1, \dots, x_n)$, $x_i \sim N(\mu, \sigma^2)$

$\hookrightarrow$ conocida

- función de verosimilitud

$$L(\theta|\underline{x}) = \prod_{i=1}^{n} f(x_i|\mu,\sigma^2) \propto \prod_{i=1}^{n} \exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}\left(n\mu^2 - 2n\mu\bar{x}\right)\right\}$$

- apriori $\longrightarrow$ aposteriori

$\rightarrow$ APRIORI DE JEFFREY

$$\ln f(x|\mu) = -\frac{(x-\mu)^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi\sigma)$$

$$\frac{\partial}{\partial\mu}\ln f(x|\mu) = \frac{1}{\sigma^2}(x-\mu)$$

$$\hookrightarrow I(\mu) = E\left[\frac{1}{\sigma^2}(X-\mu)\right] = \frac{1}{\sigma^4}\cdot\sigma^2 = 1/\sigma^2$$

$\Rightarrow \pi(\mu) = 1/\sigma \rightsquigarrow$ COSTANTE: coincide con una no informativa compl.

$$\Rightarrow \pi(\mu|\underline{x}) \propto \exp\left\{-\frac{1}{2\sigma^2}\left(n\mu^2 - 2n\mu\bar{x}\right)\right\}$$

$$\Rightarrow \mu|\underline{x} \sim N\left(\bar{x}, \frac{\sigma^2}{n}\right)$$

→ APRIORI CON EL METODO DE LA CONJUGADAS

Está claro que la clase de distribuciones conjugadas a la función de verosimilitud es la de las normales

Por esto, $\mu \sim N(\mu_0, \sigma_0^2)$

$$\Rightarrow \pi(\mu|\underline{x}) \propto \exp\left\{-\frac{1}{2\sigma^2}\left(n\mu^2 - 2n\mu\bar{x}\right)\right\}\exp\left\{-\frac{1}{2\sigma_0^2}\left(\mu^2 - 2\mu_0\mu\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\mu^2\left(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}\right) - 2\mu\left(\frac{n}{\sigma^2}\bar{x} + \frac{1}{\sigma_0^2}\mu_0\right)\right)\right\}$$

$$\Rightarrow \mu|\underline{x} \sim N\left(\frac{n\sigma_0^2\bar{x} + \sigma^2\mu}{n\sigma_0^2 + \sigma^2}, \frac{\sigma_0^2\sigma^2}{n\sigma_0^2 + \sigma^2}\right)$$

- estimadores y intervalos de credibilidad

Asumiendo como aposteriori $\mu|\underline{x} \sim N(\mu_p, \sigma_p^2)$

$\hat{\mu}_B = \mu_p$ (para las funciones de perdida cuadr. y L^1, Normal es simetrica)

Intervalo de credibilidad

$$IC(\mu, 1-\alpha) = \mu_p \pm z_{1-\frac{\alpha}{2}} \cdot \sigma_p$$

- distribución predictiva

$Y \sim N(\mu, \sigma^2)$ → nueva obs. iid

$$\hat{f}(y|\underline{x}) = \int_{-\infty}^{+\infty} f(y|\mu,\underline{x})\,\pi(\mu|\underline{x})\,d\mu$$

para calcularlo usamo la función generatriz de momentos

$$m_{y|\underline{x}}(t) = \int_{-\infty}^{+\infty} e^{ty}\,\hat{f}(y|\underline{x})\,dy$$

$$= \int_{-\infty}^{+\infty} e^{ty}\left[\int_{-\infty}^{+\infty} \hat{f}(y|\mu,\underline{x})\,\pi(\mu|\underline{x})\,d\mu\right]dy$$

$$\overset{\text{FUBINI TONELLI}}{=} \int_{-\infty}^{+\infty} \pi(\mu|\underline{x})\left[\int_{-\infty}^{+\infty} e^{ty}\,f(y|\mu,\underline{x})\,dy\right]d\mu$$

$$= \int_{-\infty}^{+\infty} \pi(\mu|\underline{x})\,m_{y|\mu}(t)\,d\mu$$

$$= \int_{-\infty}^{+\infty} \pi(\mu|\underline{x})\exp\left\{\mu t + \frac{t^2\sigma^2}{2}\right\}d\mu$$

$$= e^{\frac{t^2\sigma^2}{2}}\int_{-\infty}^{+\infty} \pi(\mu|\underline{x})\,e^{\mu t}\,d\mu \quad = m_{\mu|\underline{x}}(t)$$

$$= e^{t^2\sigma^2/2}\cdot e^{\mu_p t + \frac{1}{2}\sigma_p^2 t^2} = m_{y|\underline{x}}(t)$$

$$\Rightarrow \boxed{y \mid \underline{x} \sim \mathcal{N}(\mu_p,\ \sigma^2 + \sigma_p^2)}$$

MODELO NORMAL (varianza desconocida)

- función de verosimilitud

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

$$L(\mu, \sigma^2 \mid \underline{x}) = (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \underbrace{\sum_{i=1}^{n} (x_i - \mu)^2}_{} \right\}$$

$$= \sum_{i=1}^{n} (x_i - \bar{x} + \bar{x} - \mu)^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + \cancel{2(\bar{x}-\mu)\sum_{i=1}^{n}(x_i - \bar{x})} + n(\bar{x}-\mu)^2$$

notación :

$$\mu \longleftrightarrow \theta$$

$$\sigma^2 \longleftrightarrow \phi$$

$$L(\theta, \phi \mid \underline{x}) = (2\pi\phi)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\phi} \left((n-1)S^2 + n(\bar{x} - \theta)^2 \right) \right\}$$

- apriori $\longrightarrow$ aposteriori

$$I(\theta, \phi) = \begin{pmatrix} -E\left[\frac{\partial^2}{\partial\theta^2} \log f(X \mid \theta, \phi) \right] & -E\left[\frac{\partial^2}{\partial\theta\partial\phi} \log f(X \mid \theta, \phi) \right] \\ -E\left[\frac{\partial^2}{\partial\theta\partial\phi} \log f(X \mid \theta, \phi) \right] & -E\left[\frac{\partial^2}{\partial\phi^2} \log f(X \mid \theta, \phi) \right] \end{pmatrix}$$

$$\left(\log f(x \mid \theta, \phi) = -\frac{(x-\theta)^2}{2\phi} - \frac{1}{2} \log(2\pi\phi) \right)$$

APRIORI DE JEFFREY

- $\frac{\partial^2}{\partial\theta^2} \log f(X \mid \theta, \phi) = \frac{\partial}{\partial\theta}\left[\frac{(x-\theta)}{\phi} \right] = -1/\phi$

- $\frac{\partial^2}{\partial\theta\partial\phi} \log f(X \mid \theta, \phi) = \frac{\partial}{\partial\phi}\left[\frac{(x-\theta)}{\phi} \right] = -\frac{(x-\theta)}{\phi^2}$

$$\bullet\ \frac{\partial^2}{\partial\phi^2}\log f(x\mid\theta,\phi)=\frac{\partial}{\partial\phi}\left[\frac{(x-\theta)^2}{2\phi^2}-\frac{\pi}{\phi}\right]=-\frac{(x-\theta)^2}{\phi^3}+\frac{\pi}{\phi^2}$$

$$=\begin{pmatrix}1/\phi & 0\\ 0 & \frac{1}{\phi^2}(\pi-1)\end{pmatrix}$$

$$\Rightarrow \pi(\theta,\phi)\propto\sqrt{\det I(\theta,\phi)}$$

$$\propto \phi^{-3/2}$$

$$\Rightarrow \pi(\theta,\phi\mid\underline{x})\propto\phi^{-\frac{1}{2}(n+3)}\exp\left\{-\frac{1}{2\phi}(n-1)S^2\right\}\exp\left\{-\frac{n(\theta-\bar{x})^2}{2\phi}\right\}$$

$$\propto \phi^{-(\frac{n}{2}+1)}\exp\left\{-\frac{1}{2\phi}(n-1)S^2\right\}\phi^{-\frac{1}{2}}\exp\left\{-\frac{n(\theta-\bar{x})^2}{2\phi}\right\}$$

$$\bullet\ \theta\mid\underline{x}\sim N(\bar{x},\phi/n)$$

$$\bullet\ \phi\mid\underline{x},\sim GaI\left(\frac{n}{2},\frac{(n-1)S^2}{2}\right)$$

<u>GAMMA-INVERTIDA</u> $X\sim GaI(a,\lambda)$

$$f(x\mid a,\lambda)=\frac{\lambda^a}{\Gamma(a)}\cdot\frac{1}{x^{a+1}}e^{-\lambda/x}\quad \forall x>0,\ a>0,\lambda>0$$

<u>prop</u>: $X\sim GaI(a,\lambda)$

$$\Rightarrow \frac{1}{X}\sim Ga(a,\lambda)$$

$$\Rightarrow E[X]=\frac{\lambda}{a-1},\ Var(X)=\frac{\lambda^2}{(a-1)^2(a-2)}$$

$$moda(X)=\frac{\lambda}{a+1}$$

$$\hookrightarrow (\theta,\phi)\mid\underline{x}\sim NGaI\left(\bar{x},\frac{1}{n},\frac{n}{2},\frac{(n-1)S^2}{2}\right)$$

Prop: MARGINALES

$$\rightarrow \boxed{\Phi \mid \underline{x} \sim GaI\left(\frac{n}{2}, \frac{(n-1)S^2}{2}\right)}$$

$$\Rightarrow \frac{(n-1)S^2}{\Phi} \mid \underline{x} \sim \chi^2_{(n)}$$

$$\rightarrow \pi(\theta \mid \underline{x}) = \int_0^{+\infty} \pi(\theta, \Phi \mid \underline{x})\, d\Phi$$

$$\lambda^* = \frac{1}{2}\left[n(\theta - \bar{x})^2 + (n-1)S^2\right]$$

$$\propto \underbrace{\int_0^{+\infty} \frac{1}{\Phi^{\frac{n+1}{2}+1}} \exp\left\{-\frac{1}{\Phi}\lambda^*\right\} d\Phi}_{\text{GAMMA - INVERTIDA}}$$

$$\propto \left(1 + \frac{z^2}{n}\right)^{-\frac{n+1}{2}} \quad \forall z \in \mathbb{R} \text{ donde } z = \sqrt{n}\,\frac{(\theta - \bar{x})}{S}$$

$$\Rightarrow z \mid \underline{x} \sim t_{(n)}$$

→ APRIORI CON EL METODO DE LAS CONJUGADAS

apriori $(\theta, \Phi) \sim NGaI(m, c, a, b)$

$$\pi(\theta, \Phi \mid \underline{x}) \propto \Phi^{-\frac{n}{2}} \exp\left\{-\frac{(n-1)S^2}{2\Phi}\right\} \exp\left\{-\frac{n(\theta - \bar{x})^2}{2\Phi}\right\}$$

$$\cdot\, \Phi^{-(a+1)} \exp\left\{-\frac{b}{\Phi}\right\} \cdot \Phi^{-\frac{1}{2}} \exp\left\{-\frac{(\theta - m)^2}{2c}\right\}$$

$$\propto \Phi^{-\left(a+\frac{n}{2}+1\right)} \exp\left\{-\frac{(n-1)S^2+2b}{2\Phi}\right\} \Phi^{-\frac{n}{2}} \cdot$$

$$\exp\left\{-\left(\vartheta^2\left(\frac{n}{2\Phi}+\frac{1}{2c}\right)-2\vartheta\left(\frac{n\bar{x}}{2\Phi}+\frac{m}{2c}\right)\right)\right\}$$

$$\Rightarrow (\vartheta, \Phi) \mid \underline{x} \sim NGaI\left(\frac{nc\bar{x}+m}{nc+1}, \ldots\right)$$

3) INFERENCIA PARA MODELOS MULTIPARAMETR.

$X_1, \dots, X_m$ v.a. t.q X_i : frecuencia de la categ. i

↳ t.q. total de observaciones fijo, $N = \sum_{i=1}^{m} x_i$

probabilidades para categorías: $\underline{\theta} = (\theta_1, \dots, \theta_m)$, $\theta_i \in (0,1)$

$$\sum_i \theta_i = 1$$

distr. de prob. de $(X_1, \dots, X_m)$:

$$\mathbb{P}(X_1 = x_1, \dots, X_m = x_m \mid \underline{\theta}) = \frac{N!}{\prod_{i=1} x_i!} \prod_{i=1}^{m} \theta_i^{x_i}$$

obs : $m = 2$: $\mathbb{P}(X_1 = x_1, X_2 = x_2 \mid \theta_1, \theta_2) = \frac{(x_1+x_2)!}{x_1!\, x_2!} \theta_1^{x_1} \theta_2^{x_2}$

↳ Beta - Binomial

apriori de Jeffrey:

$$f(\underline{x} \mid \underline{\theta}) \longrightarrow \ln f(\underline{X} \mid \underline{\theta}) = \ln(N!) - \sum_{i=1}^{m} \ln(x_i!) + \sum_{i=1}^{m} x_i \ln \theta_i$$

$$\longrightarrow \frac{\partial}{\partial \theta_i} \ln f(\underline{X} \mid \underline{\theta}) = \frac{x_i}{\theta_i} \quad , \quad \frac{\partial^2}{\partial \theta_i^2} \ln f(\underline{X} \mid \underline{\theta}) = -\frac{x_i}{\theta_i^2}$$

$$\longrightarrow \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln f(\underline{X} \mid \underline{\theta}) = 0$$

$X_i \sim Bin(N, \theta_i)$

$$\Rightarrow \mathbb{E}\left[-\frac{\partial^2}{\partial \theta_i^2} \ln f(\underline{X} \mid \underline{\theta}) \right] = \mathbb{E}\left[\frac{X_i}{\theta_i^2} \right] = \frac{N}{\theta_i}$$

$$\Rightarrow \det I(\theta_1, \dots, \theta_m) = \prod_{i=1}^{m} \frac{N}{\theta_i}$$

$$\Rightarrow \pi(\theta_1, \dots, \theta_m) \propto \prod_{i=1}^{m} \theta_i^{-1/2}$$

es decir $\underline{\theta} \sim Di\left(\frac{1}{2}, \dots, \frac{1}{2}\right)$

apriori conjugada : $\underline{\theta} \sim Di(a_1, \dots, a_m)$

De hecho, la aposteriori se calcula como

$$\pi(\underline{\theta} \mid \underline{x}) \propto \prod_{i=1}^{m} \theta_i^{x_i} \cdot \prod_{i=1}^{m} \theta_i^{a_i - 1}$$

$$\Rightarrow \underline{\theta} \mid \underline{x} \sim Di(x_1 + a_1, \dots, x_m + a_m)$$

TABLAS DE CONTINGENCIA

Sin perdida de generalidad estudiamos el caso de tablas 2x2

Dos variables cualitativas (X, Y), cada una con dos modalidades

→ desde una muestra de tamaño n se observan las frecuencias de cada caso

n_{ij} : nº obs. de x_i conjunta a y_j $\forall i,j = 1,2$

$$\theta_{ij} := \mathbb{P}(X = x_i, Y = y_j) \quad \forall i,j$$

objetivo: Establecer si existe independencia entre X e Y

$$\begin{cases} H_0 : M_D \longrightarrow \text{no independencia} \\ H_1 : M_I \longrightarrow \text{independencia} \end{cases}$$

$$\hookrightarrow FB_{01} = \frac{\mathbb{P}(\underline{x} \mid M_D)}{\mathbb{P}(\underline{x} \mid M_I)}$$

tabla de contingencia:

	y_1	y_2
x_1	θ_{11}	θ_{12}
x_2	θ_{21}	θ_{22}

$$\Theta = \left\{ (\theta_{ij})_{ij} \text{ t.q. } \theta_{ij} > 0 \ \Sigma\theta_{ij} = 1 \right\}$$

apriori: $\underline{\theta} \sim Di(1,1,1,1)$

(1) MODELO NO-INDEPENDENCIA

$$\mathbb{P}(\underline{x} \mid M_D) = \int_{\Theta} \mathbb{P}(\underline{x} \mid \underline{\theta}, M_D)\, \mathbb{P}(\underline{\theta} \mid M_D)\, d\underline{\theta}$$

$$= \frac{n!}{\prod n_{ij}!} \prod_{ij} \theta_{ij}^{n_{ij}} \quad \text{donde } \underline{x} = (n_{11}, n_{12}, n_{21}, n_{22})$$

$$= \int_{\Theta} \frac{n!}{\prod n_{ij}!} \prod_{ij} \theta_{ij}^{n_{ij}} \frac{1}{Di(1,1,1,1)}\, d\underline{\theta} = \frac{n!}{\prod n_{ij}!} \frac{Di(n_{11}+1, \dots, n_{22}+1)}{Di(1,1,1,1)}$$

(2) MODELO DE INDEPENDENCIA

restricción de indep. : $\theta_{ij} = \theta_{i+}\,\theta_{j+} \quad \forall i,j$

$$\left(\theta_{i+} = \theta_{i1} + \theta_{i2} \right)$$

$$\Rightarrow \mathbb{P}(\underline{x} \mid \underline{\theta}, M_I) = \frac{n!}{\prod n_{ij}!}\, \theta_{1+}^{n_{1+}}\, \theta_{2+}^{n_{2+}}\, \theta_{+1}^{n_{+1}}\, \theta_{+2}^{n_{+2}}$$

priori se hace : $\mathbb{P}(\underline{\theta}_+ \mid M_I) = \dfrac{1}{D(1,1)} \cdot \dfrac{1}{D(1,1)}\, I_{\Theta_+}(\underline{\theta}_+)$

$$\Rightarrow \mathbb{P}(\underline{x} \mid M_D) = \frac{n!}{\prod n_{ij}!} \cdot \frac{D_i(n_{1+}+1,\, n_{2+}+1)\, D(n_{+1}+1,\, n_{+2}+1)}{D_i(1,1)\, D_i(1,1)}$$