Recordar

$$X_1 \dots X_n \overset{iid}{\sim} F_X(x;\theta) \qquad \theta \in \Theta \subset \mathbb{R}^k$$

$$L_n(\theta) = f_{X_1 \dots X_n}(x_1 \dots x_n; \theta) = \prod_{i=1}^{n} f_X(x_i;\theta)$$

$$\hat{\theta}^{MV} = \underset{\theta \in \Theta}{\operatorname{argmax}} \; L_n(\theta)$$

Usualmente existe y es único.

También definimos $\ell_n(\theta) = \operatorname{Log} L_n(\theta)$.

Con regularidad $\hat{\theta}^{MV}$ se obtiene resolviendo

$$\underbrace{\nabla \ell_n(\theta)}_{\text{función de Score}} = 0 \qquad (\text{sist } k \text{ ecuaciones})$$

<u>Def</u>

$$I(\theta)_{ij} = E\left\{\left[\frac{\partial}{\partial \theta_i} \ell_n(\theta)\right]\left[\frac{\partial}{\partial \theta_j} \ell_n(\theta)\right]\right\}$$

con reg. $$= -E\left\{\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ell_n(\theta)\right\}$$

Con regularidad

$$\sqrt{n}\left[\hat{\theta}^{MV} - \theta\right] \overset{\mathcal{D}}{\Longrightarrow} N_k\left(0; I^{-1}(\theta)\right)$$

Esto implica que para n suff gde

$$\sqrt{n}\left[\hat{\theta}^{MV} - \theta\right] \overset{A}{\sim} N_k\left(0; \underbrace{I^{-1}(\theta)}_{\text{depende parámetro}}\right)$$

$$I^{1/2}(\theta)\sqrt{n}\left[\hat{\theta}^{MV} - \theta\right] \overset{A}{\sim} N_k\left(0, I\right)$$

identidad

Hecho también vale si remplazo $I(\theta)$ por $I(\hat{\theta}^{MV})$

o sea

$$I^{1/2}(\hat{\theta}^{MV})\sqrt{n}\left[\hat{\theta}^{MV}-\theta\right] \overset{A}{\sim} N_k(0, I)$$

$$\therefore \quad \boxed{\sqrt{n}\left[\hat{\theta}^{MV}-\theta\right] \overset{A}{\sim} N_k\left(0; I^{-1}(\hat{\theta}^{MV})\right)}$$

Fórmula utilizable.

Notar nuestra deducción se basó en que

1) $$\ell_n(\theta) = \underbrace{\sum_{i=1}^{n} \log f_X(X_i;\theta)}_{\text{Terminos iid} \cdots \text{TCL, LGN}}$$

2) $$\ell_n'(\hat{\theta}^{MV}) \cong \ell_n'(\theta) + \underbrace{\ell_n''(\theta)}_{0}(\hat{\theta}^{MV}-\theta)$$

Observaciones si $X_1 \dots X_n$ no son iid, ej serie tiempo el paso 1 no vale. Frecuentemente hay teoremas análogos para estos casos.

$X_1 \dots X_n$ es "m-dependiente" ssi

$$\operatorname{cov}(X_n; X_{n+m+k}) = 0 \quad \forall n, k \in \mathbb{N}$$

¿Qué puede fallar?

1) Cuando no es solución ec. score
2) Que el resto de taylor no sea despreciable
3) Que el término segundo en taylor sea cero.

etc

¿

Método Delta

Situación

tenemos $\sqrt{n}\,(\tilde{\theta}-\theta) \Rightarrow N(0, \Sigma)$

nos interesa $g(\theta)$ $\quad g: \mathbb{R}^k \to \mathbb{R}^q \quad q \leq k$

continuo y diferenciable

queremos

$$\sqrt{n}\,(g(\tilde{\theta}) - g(\theta)) \Rightarrow \;?$$

Desarrollo

$$g(\tilde{\theta}) \cong g(\theta) + \nabla g(\theta)'(\tilde{\theta}-\theta)$$

$$\sqrt{n}\,[g(\tilde{\theta}) - g(\theta)] \cong \nabla g(\theta)' \underbrace{\sqrt{n}\,[(\tilde{\theta}-\theta)]}_{N_k(0,\, \Sigma)}$$

$$\boxed{\sqrt{n}\,[\underset{(q\times 1)}{g(\tilde{\theta})} - \underset{(q\times 1)}{g(\theta)}] \Rightarrow N_q\Big(\underset{(q\times 1)}{0}\,;\ \underbrace{\underset{q\times k}{\nabla' g(\theta)}\ \underset{k\times k}{\Sigma}\ \underset{k\times q}{\nabla g(\theta)}}_{q\times q}\Big)}$$

$$g(\theta) = \begin{pmatrix} g_1(\theta_1, \ldots, \theta_k) \\ \vdots \\ g_q(\theta_1, \ldots, \theta_k) \end{pmatrix}$$

$$\nabla' a = \nabla a'$$

$$\underset{(q\times k)}{Dg(\theta)} = \begin{pmatrix} \dfrac{\partial g_1}{\partial \theta_1} & \cdots & \dfrac{\partial g_1}{\partial \theta_k} \\ \vdots & & \\ \dfrac{\partial g_q}{\partial \theta_1} & \cdots & \dfrac{\partial g_q}{\partial \theta_k} \end{pmatrix} = \nabla' g(\theta)$$

Observar $\sqrt{n}(\tilde{\theta}-\theta) \Rightarrow N(0, \sigma^2)$

implica $\tilde{\theta}$ es consistente

$$\frac{1}{\cancel{\sqrt{n}}}\{\cancel{\sqrt{n}}(\tilde{\theta}-\theta)\} \Rightarrow \frac{1}{\sqrt{n}} N(0, \sigma^2)$$

$$\tilde{\theta}-\theta \Rightarrow \text{masa puntual en } 0$$

$$\tilde{\theta}-\theta = 0 \text{ con probabilidad } 1$$

$$\tilde{\theta} \xrightarrow{P} \theta.$$

también valido si

$$n^{\alpha}(\tilde{\theta}-\theta) \Rightarrow F$$

Ej $X_1 \ldots X_n$ iid $B(\theta)$ $\qquad \hat{\theta} = \frac{1}{n}\sum X_i$

$$\sqrt{n}(\hat{\theta}-\theta) \Rightarrow N(0, \theta(1-\theta))$$

$$\sqrt{n}(\hat{\theta}-\theta) \overset{A}{\sim} N(0, \hat{\theta}(1-\hat{\theta}))$$

IC $(1-\alpha)100\%$.

$$\hat{\theta} \pm z_{1-\alpha/2} \frac{1}{\sqrt{n}}\sqrt{\hat{\theta}(1-\hat{\theta})} \qquad \text{usual}$$

Otra opción. Partir de $\sqrt{n}(\hat{\theta}-\theta) \Rightarrow N(0, \theta(1-\theta))$

encontrar $g(\theta)$ tal que $\sqrt{n}(g(\hat{\theta})-g(\theta)) \Rightarrow N(0,1)$

Por delta

$$\sqrt{n}(g(\hat{\theta})-g(\theta)) \Rightarrow N(0; [g'(\theta)]^2\theta(1-\theta))$$

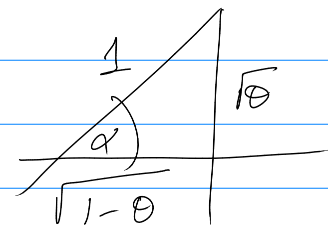
debo resolver

$$[g'(\theta)]^2\theta(1-\theta) = 1$$

$$\left[\frac{d}{d\theta} g(\theta)\right]^2 = \frac{1}{\theta(1-\theta)} \qquad \frac{d}{d\theta} g(\theta) = \frac{1}{\sqrt{\theta(1-\theta)}}$$

$$\int d\, g(\theta) = \int \frac{1}{\sqrt{\theta(1-\theta)}}\, d\theta$$

$$g(\theta) = \int \frac{1}{\sqrt{\theta(1-\theta)}}\, d\theta$$



$$\sqrt{\theta} = \sin\alpha$$
$$\sqrt{1-\theta} = \cos\alpha \qquad \text{etc}$$
$$\theta = \sin^2\alpha$$

$$g(\theta) = \arcsin\theta$$

$$\sqrt{n}\left[\arcsin\hat{\theta} - \arcsin\theta\right] \Rightarrow N(0,1)$$

IC $(1-\alpha)100\%$ para arcsin

$$\arcsin\hat{\theta} \pm z_{1-\alpha/2} \frac{1}{\sqrt{n}}$$

y para volver a θ

IC $(1-\alpha)100\%$

$$\left[\sin\left(\arcsin\hat{\theta} - z_{1-\alpha/2}\frac{1}{\sqrt{n}}\right); \sin\left(\arcsin\hat{\theta} + z_{1-\alpha/2}\frac{1}{\sqrt{n}}\right)\right]$$

En este caso la transformación arcsin θ se llama transformación estabilizadora de varianza
Simulación muestra que mejora la aproximación.

Ej $X_1 \dots X_n \overset{iid}{\sim} P(\lambda) \qquad \hat{\lambda} = \frac{1}{n}\sum X_i$

$$\boxed{\sqrt{n}\,(\hat{\lambda}^{mv} - \lambda) \overset{D}{\Rightarrow} N(0, \lambda)}$$

$$\sqrt{n}\,[g(\hat{\lambda}^{mv}) - g(\lambda)] \overset{D}{\Rightarrow} N(0,1)$$

$$\Rightarrow \boxed{N(0, [g'(\lambda)]^2 \lambda)}$$

$$g'(\lambda)^2 \lambda = 1 \qquad g'(\lambda) = \frac{1}{2\sqrt{\lambda}} \quad 2$$

$$g(\lambda) = 2\sqrt{\lambda} \qquad \sqrt{n}\,(2\sqrt{\hat{\lambda}} - 2\sqrt{\lambda}) \Rightarrow N(0,1)$$

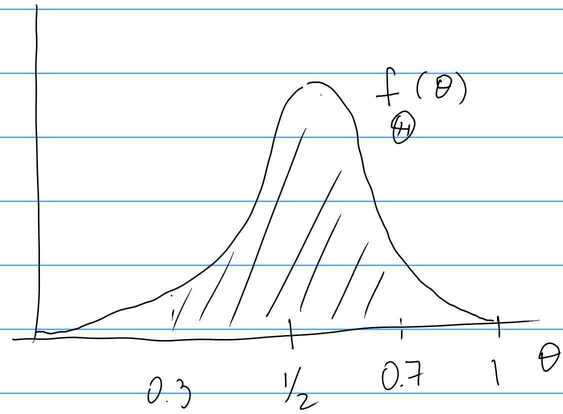
y en este caso los dos intervalos competidores son

usual $\qquad \hat{\lambda} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\lambda}}{n}}$

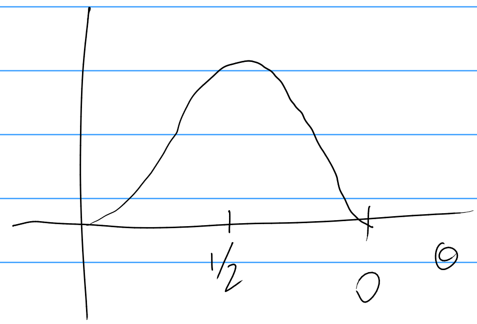
$2\sqrt{\lambda} \qquad \left(2\sqrt{\hat{\lambda}} - z_{1-\alpha/2}\frac{1}{\sqrt{n}} \,;\, 2\sqrt{\hat{\lambda}} + z_{1-\alpha/2}\frac{1}{\sqrt{n}}\right)$

$$\left[\left(\sqrt{\hat{\lambda}} - \frac{1}{2} z_{1-\alpha/2}\frac{1}{\sqrt{n}}\right)^2 ; \left(\sqrt{\hat{\lambda}} + \frac{1}{2} z_{1-\alpha/2}\frac{1}{\sqrt{n}}\right)^2\right]$$

Estadística Bayesiana



Estadística Clásica $\rightarrow$ Probabilidad Frecuentista ✓
Estadística Bayesiana $\rightarrow$ Probabilidad Subjetiva ✓



$P: \mathcal{a} \rightarrow [0,1]$ $\quad (\Omega, \mathcal{a}, P)$

i) $P(\Omega) = 1$

ii) $P(A^c) = 1 - P(A)$

iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

$A_i \cap A_j = \emptyset$

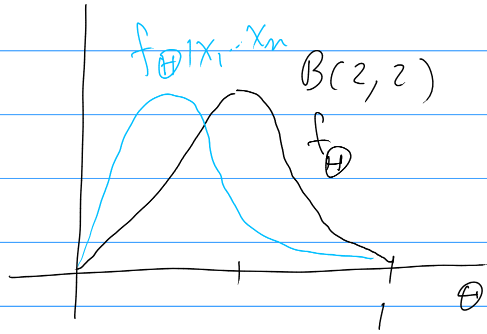
En estadística Bayesiana Θ es una v.a. con distribución apriori $F_{\Theta}(\theta)$ que refleja lo que se sabe o cree de un parámetro antes de tomar ninguna muestra.

Ej $\quad \Theta \sim \text{Beta}(2,2)$

v.a. $\quad$ realización

$X \rightarrow x$

$\Theta \rightarrow \theta$



$$f_{\Theta}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$E\,\Theta = \dfrac{\alpha}{\alpha+\beta}$ $\qquad$ Var ... fn α, β

$X_1 \ldots X_n$ muestra aleatoria $X|\Theta=\theta \sim B(\theta)$

$$f_{X_1\ldots X_n|\Theta=\theta}(x_1\ldots x_n;\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

distr condicional de la muestra dado el parametro. verosimilitud Modelo.

$$f_{\Theta|X_1=x_1,\ldots X_n=x_n}(\theta; x_1, x_2, \ldots x_n) = \frac{f_{X_1\ldots X_n;\Theta}(x_1\ldots x_n;\theta)}{\int f_{X_1\ldots X_n;\Theta}(x_1\ldots x_n,\theta)\,d\theta}$$

distribucion a posteriori

$$f_{\Theta}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$f_{X_1\ldots X_n;\Theta}(x_1\ldots x_n;\theta) = f_{\Theta}(\theta)\, f_{X_1\ldots X_n|\Theta=\theta}(x_1\ldots x_n;\theta)$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\cdot \underbrace{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}}_{\binom{n}{t}\theta^{t}(1-\theta)^{n-t}} = f_{T,\Theta}(t,\theta)$$

Nota si mi muestra solo tuviera $T=\sum X_i$

$T \sim B(n,\theta)$

$$f_{X_1\ldots X_n}(x_1\ldots x_n) = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha+\sum x_i-1}(1-\theta)^{\beta-1+n-\sum x_i}\,d\theta$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+\sum x_i)\Gamma(\beta+n-\sum x_i)}{\Gamma(\alpha+\beta+n)}\underbrace{\frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+\sum x_i)\Gamma(\beta+n-\sum x_i)}\int_0^1\theta^{\alpha+\sum x_i-1}(1-\theta)^{\beta+n-\sum x_i-1}\,d\theta}_{1}$$

$$f_{X_1 \dots X_n}(x_1 \dots x_n) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+\sum x_i)\Gamma(\beta+n-\sum x_i)}{\Gamma(\alpha+\beta+n)}$$

$$f_{\Theta|X_1, X_2 \dots X_n}(\theta, x_1 \dots x_n) = \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+\sum x_i)\Gamma(\beta+n-\sum x_i)} \theta^{\alpha+\sum x_i - 1} (1-\theta)^{\beta+n-\sum x_i - 1}$$

$$\Theta | X_1 = x_1 \dots X_n = x_n \sim \text{Beta}\left(\underbrace{\alpha + \sum x_i}\ ;\ \underbrace{\beta + n - \sum x_i}\right)$$

α y β se llaman "hiperparámetros".

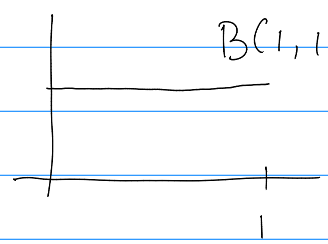
Ahora necesitamos resumir la distribución a posteriori.

Una opción

$$\hat{\theta}^B = E\left[\Theta | X_1 = x_1 \dots X_n = x_n\right]$$

$$= \frac{\alpha + \sum x_i}{\alpha + \cancel{\sum x_i} + \beta + n - \cancel{\sum x_i}}$$

$$= \frac{\sum x_i + \alpha}{n + \alpha + \beta}$$



Notar

$$\hat{\theta}^B = \underbrace{\frac{n}{n+\alpha+\beta}}_{\substack{\to 1 \\ n\to\infty}} \underbrace{\frac{\sum x_i}{n}}_{\hat{\theta}^{MV}} + \underbrace{\frac{\alpha+\beta}{n+\alpha+\beta}}_{\substack{\to 0 \\ n\to\infty}} \underbrace{\frac{\alpha}{\alpha+\beta}}_{E(\text{priori})}$$

MAL

$$\left| \hat{\theta}^B - \hat{\theta}^{MV} \right| = o(1)$$

$$\Theta \sim U(1/4, 3/4)$$

La priori puede provenir de

1) Experiencia pasada / similar
2) Conocimientos teoricos GINI
3) Conveniencia matematica

Gibbs
