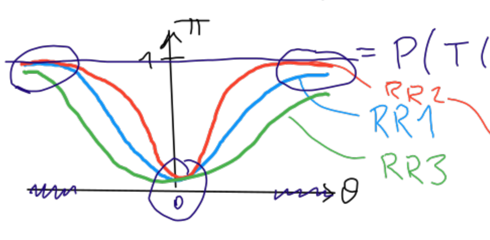
Potencia: $\pi(\theta) = P(\text{rechazar } H_0 \; \forall \theta)$

$= P(T(\underline{X}) \in RR \; \forall \theta)$



la función potencia más "potente" de la 3

__Ejemplo__ Consideremos una población Ber(p

Queremos contrastar $H_0: p \leq \frac{1}{2}$ vs $\boxed{H_1: p > \frac{1}{2}}$

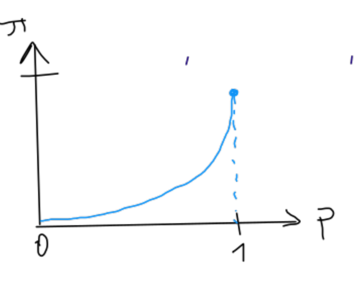
Suponer que se obtiene una muestra $X_1, \dots, X_5$

a) Consideremos el contraste de hip. que rechaza H_0 si en la muestra sólo se obtienen éxitos. Encontrar $\pi(p)$

$$RR = \left\{ \underline{X} : \sum_{i=1}^{5} X_i = 5 \right\}$$

0 1 2 3 4 5 $\quad \sum X_i = 5$

$$\pi_1(p) = P\Big(\underbrace{\sum X_i}_{\sim Bin(5,p)} = 5\Big) = \frac{5!}{5!\,0!} p^5 (1-p)^0 = p^5$$

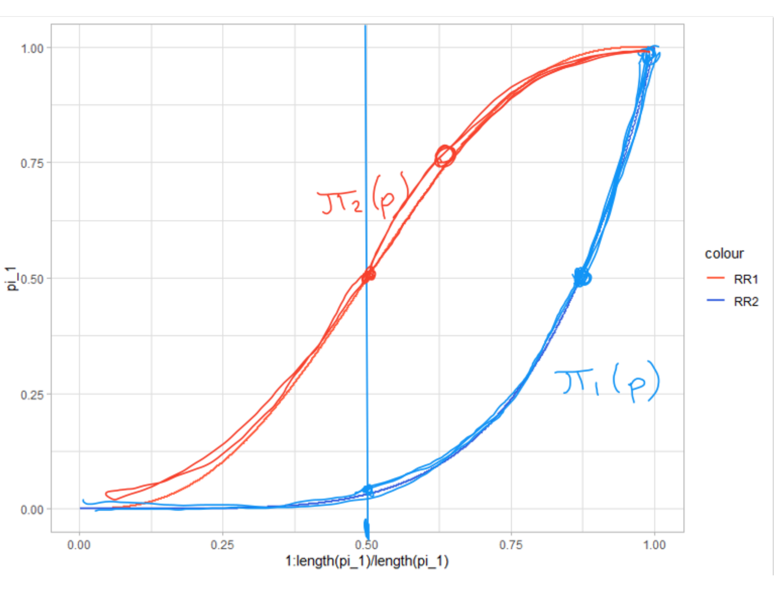


b) Considerar el contraste de hip. que rechaza H_0 si $\sum X_i = 3, 4, 5$. Encontrar $\pi(p)$

$$RR = \left\{ \underline{X} : \sum_{i=1}^{5} X_i \geq 3 \right\}$$

0 1 2 3 4 5 $\quad \sum X_i \geq 3$

$$\pi_2(p) = P\left(\sum X_i \geq 3\right) = \frac{5!}{3!\,2!} p^3 (1-p)^2 + \frac{5!}{4!\,1!} p^4 (1-p)^1 + \frac{5!}{5!\,0!} p^5 (1-p)^0$$



$\pi_2(p) \geq \pi_1(p)$

$\forall \; p \in (0,1)$

$\therefore \pi_2(p)$ es más potente

Lema de Neyman-Pearson

$H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$

(simple vs simple)

La RR asociada a la prueba mas potente es

$$RR = \left\{ \underline{X} : \frac{\mathcal{L}(\theta_0; \underline{x})}{\mathcal{L}(\theta_1; \underline{x})} \leq k_\alpha \right\}$$

donde α es el nivel de significancia

__Ejemplo__ $X_1, \dots, X_n$ m.a. de

$$f_X(x; \theta) = \theta x^{\theta - 1} I_{(0,1)}(x), \quad \theta > 0$$

Queremos probar $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$, $\theta_0 < \theta_1$

a) Probar que la RR asociada a la prueba más potente es $RR = \left\{ \underline{X} : \boxed{-\sum_{i=1}^{n} \log(X_i)} \leq k_\alpha^* \right\}$

Por el lema de NP, tenemos que

$$RR = \left\{ \underline{X} : \frac{\mathcal{L}(\theta_0; \underline{x})}{\mathcal{L}(\theta_1; \underline{x})} \leq k_\alpha \right\}$$

Primero, calculemos $\mathcal{L}(\theta; \underline{x})$

$$\mathcal{L}(\theta; \underline{x}) = \prod_{i=1}^{n} \theta x_i^{\theta-1} I_{(0,1)}(x_i)$$

$$= \theta^n \prod_{i=1}^{n} e^{\log(x_i^{\theta-1})}$$

no hace nada

$$= \theta^n e^{\sum \log(x_i^{\theta-1})}$$

$$= \theta^n e^{(\theta-1)\sum \log(x_i)} I_{\mathbb{R}^+}(\theta)$$

$$\therefore \frac{\mathcal{L}(\theta_0; \underline{x})}{\mathcal{L}(\theta_1; x)} = \frac{\theta_0^n e^{(\theta_0 - 1)\sum \log(x_i)}}{\theta_1^n e^{(\theta_1 - 1)\sum \log(x_i)}} \leq k_\alpha$$

$$\iff \left(\frac{\theta_0}{\theta_1}\right)^n e^{(\theta_0 - \theta_1)\sum \log(x_i)} \leq k_\alpha$$

$$\iff e^{(\theta_0 - \theta_1)\sum \log(x)} \leq k_\alpha' \qquad k_\alpha' = \frac{k_\alpha}{\left(\frac{\theta_0}{\theta_1}\right)^n}$$

(pero no nos importa)

$$\iff \underbrace{(\theta_0 - \theta_1)}_{<0} \sum \log(x_i) \leq k_\alpha''$$

$\theta_0 < \theta_1$

$\theta_0 - \theta_1 < 0$

$$\iff \sum \log(x_i) \geq k_\alpha'''$$

$$\iff -\sum \log(x_i) \leq k_\alpha^*$$

b) Probar que $k_\alpha^* = \dfrac{2\chi^2_{2n,\alpha}}{\theta_0}$

Ya tenemos el estadístico de prueba: $-\sum \log(X_i)$

lo que queremos ver es cómo se distribuye

Sea $Y_i = -\log(X_i)$

$$\boxed{F_{Y_i}(y)} = P(Y_i \leq y) = P(-\log(X_i) \leq y)$$

$$= P(X_i \geq e^{-y}) = 1 - P(X_i \leq e^{-y})$$

$$= 1 - F_{X_i}(e^{-y})$$

$$= 1 - \int_0^{e^{-y}} \theta x^{\theta-1}\, dx = 1 - \cancel{\theta} \frac{x^\theta}{\cancel{\theta}} \Big|_0^{e^{-y}}$$

$$= \boxed{1 - e^{-\theta y}} \quad Y_i \sim Exp(\theta)$$

$$\Rightarrow \sum Y_i = -\sum \log(X_i) \sim ga(n, \theta)$$

Recordemos las siguientes propiedades

1) $G \sim ga(\alpha, \beta) \Rightarrow cG \sim ga\left(\alpha, \frac{\beta}{c}\right)$

2) $G \sim ga\left(\frac{n}{2}, 2\right) \Rightarrow G \sim \chi^2_n$

$$\sum Y_i \sim ga(n, \theta)$$

$$\Rightarrow \frac{\theta}{2} \sum Y_i \overset{\text{Prop.1}}{\sim} ga\left(n, \frac{\theta}{\frac{\theta}{2}}\right) \equiv ga(n, 2)$$

$$\Rightarrow \frac{\theta}{2} \sum Y_i \overset{\text{Prop.2}}{\sim} \chi^2_{2n}$$

$$-\frac{\theta}{2} \sum \log(X_i) \sim \chi^2_{2n}$$

Bajo H_0: $RR = \left\{ \underline{X} : \frac{\theta_0}{2} \sum (-\log(X_i)) \leq \chi^2_{2n,\alpha} \right\}$

$$= \left\{ \underline{X} : -\sum \log(X_i) \leq \frac{2\chi^2_{2n,\alpha}}{\theta_0} \right\}$$

Ejemplo

En las clínicas del IMSSS con camas para pacientes con COVID se desea estimar el número de personas que deberán ser internadas tras realizarse la prueba; esto con el fin de garantizar el abasto de equipo médico.

Sea X_i el número de pacientes con COVID internados en la clínica i en un día. Suponer que $X_i \sim Po(\lambda)$

a) Si se tiene una m.a. $X_1, \dots, X_n$, obtener la RR asociada a la prueba más potente si

$$H_0: \lambda = \lambda_0 \quad vs \quad H_1: \lambda = \lambda_1, \quad \text{con } \lambda_0 < \lambda_1$$

$$\mathcal{L}(\lambda; \underline{x}) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} I_{\mathbb{Z}_0^+}(x_i)$$

$$= \frac{e^{-n\lambda} \lambda^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!}$$

$\therefore$ Por el lema de NP,

$$RR = \left\{ \underline{X} : \frac{\dfrac{e^{-n\lambda_0} \lambda_0^{\Sigma x_i}}{\cancel{\prod x_i!}}}{\dfrac{e^{-n\lambda_1} \lambda_1^{\Sigma x_i}}{\cancel{\prod x_i!}}} \leq k_\alpha \right\}$$

$$= \left\{ \underline{X} : e^{-n(\lambda_0 - \lambda_1)} \left(\frac{\lambda_0}{\lambda_1} \right)^{\Sigma x_i} \leq k_\alpha \right\}$$

$$= \dots = \underline{\left\{ \underline{X} : \Sigma X_i \geq k_\alpha^* \right\}}$$

(Sabemos cómo se distribuye ΣX_i)

b) Se toma una muestra de tamaño 5

Obtener la RR con un nivel de significancia lo más cercano a 5% posible si $\lambda_0 = 6.4$ pacientes/día internados.

Sabemos que $X_i \sim Po(\lambda) \quad \forall i = 1, \dots, n$

Bajo H_0: $\Rightarrow \sum_{i=1}^{n} X_i \sim Po(n\lambda) = Po(5 \cdot 6.4) = Po(32)$

$\therefore k_\alpha^*$ es el cuantil de una Poisson (32) que acumula el $(1-\alpha)$ de proba.

Llamémosle $Po_{(32, 1-\alpha)}$

Como $\alpha = 5\%$, buscamos

$Po_{(32, 0.95)} = 41$ (en $x=41$ se acumula aprox. el 95%)

$$\therefore RR = \left\{ \underline{X} : \sum_{i=1}^{5} X_i \geq 41 \right\}$$

donde $\alpha = P(\Sigma X_i \geq 41) = 0.0512$

Como la Poisson es una v.a. discreta, no podemos obtener el valor exacto 5%

c) En una fecha particular, en esas cinco clínicas se observaron 4, 5, 8, 13, 18 pacientes internados

¿Existe evidencia para concluir que $\lambda = \lambda_1 = 2\lambda_0 = 12.8$?

$$T(\underline{x}) = \Sigma x_i = 4 + 5 + 8 + 13 + 18 = 48$$

¿Caemos en la RR?

$$¿\underbrace{T(\underline{x})}_{48} > \underbrace{k_\alpha^*}_{41}? \quad ✓$$

$\therefore$ Rechazamos H_0 en favor de H_1 y $\lambda = \lambda_1 = 12.8$.