

I: ESTRUCTURAS ESTADÍSTICAS

Definición: Una estructura estadística (e.e.) es una terna $(\Omega, \mathcal{A}, P)$, donde P es una familia de probabilidades sobre $(\Omega, \mathcal{A})$.

• Una e.e. es el pto de partida de la inferencia estadística, donde la distribución de probabilidad real que rige el experimento aleatorio es desconocida, y lo va a seguir siendo. Sin embargo, el objetivo de la inferencia es determinar cuál de las distribuciones de P es la verdadera (la supondremos siempre que está incluida en P), y así tomar una decisión acerca de la dist. a partir de los datos.

Definición: Si podemos expresar P mediante un parámetro $P = \{P_\theta : \theta \in \Theta\}$, llamamos espacio de parámetros a tal Θ . Diremos que el problema es de inferencia paramétrica si $\Theta \subset \mathbb{R}^n$ para algún $n \in \mathbb{N}$ ($\dim < \infty$). Si no ($\dim \infty$) hablaremos de inferencia no paramétrica.

Definición: Un estadístico sobre $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$ es un v.a. $T: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$.

• Como función medible que es, no depende de la familia P de probabilidades sino de los σ -álgebras.
• Un estadístico es función de los datos, y sólo de los datos. No puede ignorar los parámetros $\theta \in \Theta$.

Definición: Llamaremos estructura estadística imagen por T a la e.e. $(\Omega', \mathcal{A}', \{P_\theta^T : \theta \in \Theta\})$

Definición: Se dice que P es abs. continua respecto P' , o que P está dominada por P' , $P \ll P'$, si $P'(A) = 0 \Rightarrow P(A) = 0$. Se dice que P y P' son equivalentes si $P \ll P'$ y $P' \ll P$.

Si μ es un medida σ -finita, diremos que $P \ll \mu$ si $P \ll \mu \forall P \in \mathcal{P}$.

Diremos que $P \ll P'$ si $P \ll P' \forall P \in \mathcal{P}, \forall P' \in \mathcal{P}'$.

Definición: Dado $n \in \mathbb{N}$, llamaremos estructura estadística producto de $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$ a $(\Omega^n, \mathcal{A}^n, \{P_\theta^n : \theta \in \Theta\})$.

• En general, $P - \dots$ denotará $\dots \forall P \in \mathcal{P}$.

• En referencia, los datos de un experimento pueden ser considerados como los valores cuantales que toma una v.a. Si $X: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}, \{P_\theta: \theta \in \Theta\})$, entonces la e.e. producto corresponde con la observación de n v.a. iid $X_1, \dots, X_n$ (i.e., una muestra) con la misma distribución que X .

Proposición: Sea $(\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$ una e.e. y $n \in \mathbb{N}$. Si la e.e. está dominada por una medida σ -finita μ sobre $(\Omega, \mathcal{A})$, pongamos $P_\theta = \frac{dP_\theta}{d\mu}$, entonces

1) μ^n es σ -finita

2) μ^n domina a la e.e. producto $(\Omega^n, \mathcal{A}^n, \{P_\theta^n: \theta \in \Theta\})$, y una densidad de P_θ^n rep. μ^n es

$$P_\theta^n(\omega_1, \dots, \omega_n) := \prod_{i=1}^n P_\theta(\omega_i).$$

Definición: Un estadístico $T: (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow \mathbb{R}$ se dice sumable si $\exists E_P(T) < \infty \forall P \in \mathcal{P}$.

En el caso paramétrico podemos $E_\theta(T) \stackrel{\text{def}}{=} E_{P_\theta}(T)$.

Definición: Dos estadísticos T_1, T_2 se dicen equivalentes si $\{T_1 \neq T_2\}$ es P -nulo (i.e. coinciden, lo cual con P -nulo a b.v.), i.e. si $T_1 = T_2$ P -c.s.

Diremos que T_1, T_2 son P -independientes si lo son $\forall P \in \mathcal{P}$.

• dar n datos $\equiv$ dar n v.a. iid $\equiv$ dar muestra de tamaño n .

Por ej., sea la e.e. $(\mathbb{R}^n, \mathcal{B}^n, \{N^n(\mu, \sigma^2): \mu \in \mathbb{R}, \sigma^2 > 0\})$, y X_i las proyecciones. El teorema de Fisher (3°) afirma que son independientes los estadísticos

$$\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i \quad \equiv \quad \text{media muestral}$$

$$S^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \equiv \quad \text{varianza muestral}.$$

• (datos y muestras): Supongamos que queremos extraer una característica $X(\omega)$ de un problema Ω (p.ej., el nivel de alcohol). Realizar n observaciones $\Rightarrow$ obtener n valores $(X(\omega_1), \dots, X(\omega_n))$. Esto puede verse como tener de la v.e. padre n v.a. iid (muestra):

$$X_i: \Omega^n \longrightarrow \mathbb{R}$$

$$(\omega_1, \dots, \omega_n) \longmapsto X_i(\omega_1, \dots, \omega_n) := X(\omega_i).$$

Cada dato n corresponde con el valor concreto que toman los v.a. Las n observaciones se corresponden por tanto con

$$(X_1, \dots, X_n): \Omega^n \longrightarrow \mathbb{R}^n$$

$$(\omega_1, \dots, \omega_n) \longmapsto (X_1(\omega_1), \dots, X_n(\omega_n)).$$

Definición: Sea (X_n) una sucesión de v.a. iid definidas en $(\Omega, \mathcal{A}, P)$. Para cada $n \in \mathbb{N}$ fijo, se llama función de distribución empírica (de orden n) a

$$F_n: \mathbb{R} \times \Omega \longrightarrow [0, 1]$$

$$(x, \omega) \longmapsto F_n(x, \omega) := \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{(-\infty, x]}(X_i(\omega)).$$

• Nota que $n F_n(x, \omega)$ es el n.º de v.a. tales que $X_i(\omega) \leq x$.

Proposición:

- 1) Para $\omega \in \Omega$ fijo, $F_n(\cdot, \omega): \mathbb{R} \longrightarrow [0, 1]$ es la función de distribución de la v.a. discreta $Z_\omega: (\{1, \dots, n\}, \mathcal{P}, Q \equiv \text{distribución discreta}) \longrightarrow \mathbb{R}$, $Z_\omega(i) := X_i(\omega)$ (tiene los valores $X_1(\omega), \dots, X_n(\omega)$ con probabilidad común $\frac{1}{n}$).
- 2) Para $x \in \mathbb{R}$ fijo, $F_n(x, \cdot): \Omega \longrightarrow [0, 1]$ es una v.a. de distribución $b_n(F(x))$, donde $F(x)$ es la función de distribución de las X_i .

• La LFGN prueba que $F_n(x, \cdot) \xrightarrow{n \rightarrow \infty} F(x)$ P.-c.s. Es decir, una variable que solo depende de los datos de observaciones converge a $F(x)$, descubierto desde el punto de vista estadístico. Esto justifica el objetivo de la inferencia estadística: tomar decisiones o estimar parámetros a partir de los datos. Por este resultado prueba una convergencia puntual. El siguiente resultado es mucho más fuerte, y justifica el objetivo anterior:

Teorema (Glivenko - Cantelli, o fundamental de la Estadística Matemática): Sea (X_n) una sucesión de v.a. iid definidas en $(\Omega, \mathcal{A}, P)$, y sea F_X su función de distribución (común). Entonces, para casi todo $\omega \in \Omega$,

$$\sup_{x \in \mathbb{R}} |F_n(X, \omega) - F(x)| \xrightarrow{n \rightarrow \infty} 0.$$

SUFICIENCIA, COMPLETITUD Y LIBERTAD

Definición: Consideremos la e.e. $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$.

- a) Un suceso $A \in \mathcal{A}$ se dice libre si $P_\theta(A)$ no depende de θ , i.e., si es cte $\forall \theta$.
- b) Una sub- σ -álgebra $\mathcal{B} \subset \mathcal{A}$ se dice libre si todos sus elementos son libres.
- c) Un estadístico $T: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ se dice libre si la sub- σ -álgebra $T^{-1}(\mathcal{A}')$ es libre, i.e., si P_θ^T no depende de θ .

• Esta última afirmación se debe a que $P_\theta(T^{-1}(A')) = P_\theta^T(A')$. Es decir, que en la e.e. imagen solo tenemos una probabilidad, i.e., utilizando T se "pierde" toda la información sobre θ .

Definición: Consideremos la e.e. $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$.

- a) Una sub- σ -álgebra $\mathcal{B} \subset \mathcal{A}$ se dice suficiente si para cada $A \in \mathcal{A}$ $\exists f_A \in \bigcap_{\theta \in \Theta} P_\theta(A|B)$, esto es, un representante común de las probabilidades condicionales $P_\theta(A|B)$, es decir, si $\exists f_A$ tal que

$$P_\theta(A \cap B) = \int_B f_A dP_\theta \quad \forall \theta \in \Theta.$$

- b) Un estadístico $T: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ se dice suficiente si $T^{-1}(\mathcal{A}')$ es suficiente.

• Es obvio que, si B es suficiente para $\mathcal{P}$, y $\mathcal{P}' \subset \mathcal{P} \Rightarrow B$ es sf. para $\mathcal{P}'$. El recíproco en general no es cierto.

• Otros der un anexo multipa a) par b):

Lema: Si B es suficiente, para todo estadístico X real variable $\exists f \in \bigcap_{\theta \in \Theta} E_0(X|B)$, i.e., una versión común de las esperanzas condicionales $E_0(X|B)$.

Proposición: Sea $T: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$

T suficiente $\iff$ para cada $A \in \mathcal{A}$ $\exists g_A \in \bigcap_{\theta \in \Theta} P_\theta(A|T)$, una versión común de la prob. unde. $P_\theta(A|T)$ i.e.,

$$P_\theta(A \cap T^{-1}(A')) = \int_{A'} g_A dP_\theta^T \quad \forall \theta \in \Theta \quad \forall A' \in \mathcal{A}'.$$

Proposición: Sea $T: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$.

T suficiente $\iff$ para cada estadístico real variable X $\exists g \in \bigcap_{\theta \in \Theta} E_0(X|T)$, i.e.,

$$\int_{T^{-1}(A')} X dP_\theta = \int_{A'} g dP_\theta^T \quad \forall \theta \in \Theta, \quad \forall A' \in \mathcal{A}'.$$

• Una estadístico suficiente recoge toda la información sobre el parámetro contenido en la e.e. original, i.e., al pasar a la e.e. imagen no se pierde información. Vía este tipo de estadísticos podemos reducir las observaciones sin perder información relevante.

Teorema (Halmos - Savage): Si $\mathcal{P} = \{P_\theta: \theta \in \Theta\}$ es una familia de probabilidades en $\mathcal{A}$ dominada por una medida σ -finita μ , existe una subfamilia numerable $\{P_{\theta_n}: n \in \mathbb{N}\} \subset \mathcal{P}$ que es equivalente a $\mathcal{P}$. Además, existe una probab. hda P , llamada probabilidad dominante privilegiada, equivalente a $\mathcal{P}$ y

$P \ll \mu$, y que es
$$P = \sum_{n=1}^{\infty} \frac{1}{2^n} P_{\theta_n}.$$

1. $P_\theta(A|B)$
2. $E_0(X|B)$
3. $P_\theta(A|T)$
4. $E_0(X|T)$ 3

Lemma: Sei $(\Omega, \mathcal{A}, P_0: \Omega \in \mathcal{G})$ me e.e. dominiert par me reelle σ -finite μ, γ, B me sub- σ -alge.

1) Sea $P_0 = \frac{dP_0}{d\mu}$. Si existen v.a.r. no negativos $g_0: (\mathcal{R}, \mathcal{B}) \rightarrow \mathbb{R}$

$g_h(\Omega, \mathcal{A}) \rightarrow \mathbb{R}$ μ -integrable falls que $\varphi_0 = g_0 \circ h \Rightarrow B$ suficiente.

4) Si existe θ_0 tal $P_0: \theta \in \Theta \leq P_0$, y B es suficiente $\Rightarrow \forall \theta \in \Theta$

$$\exists p_0 = p_0 - h \in \frac{dp_0}{d\mu}, \text{ con } p_0 \text{ e } h \text{ come in 1),}$$

Teoreme (de Factorizării, de Neuman-Helms-Savage): See $(\Omega, \mathcal{A}, \mathbb{P}_0: \mathcal{O} \in \mathcal{O} \mathcal{I})$ me e.e. domide
per me modide o-finti μ γ see B me sub-o-algebra.

$$B \text{ suficiente} \iff \exists \frac{V, a. \text{ no negativos}}{g_0: (\Omega, B) \rightarrow \mathbb{R} \text{ v.o.}, h: (\Omega, \mathcal{R}) \rightarrow \mathbb{R}} : g_0 \cdot h \in \frac{dP_0}{d\mu} \quad \forall \theta \in \Theta$$

Teorema (de Factorización, de NKS, version para estadísticos) : Sea $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$ un c.e. dominado por una medida σ -finita μ , sea $P_\theta = \frac{dP_\theta}{d\mu}$ para cada θ y sea $T: (\Omega, \mathcal{A}, \{P_\theta\}) \rightarrow (\Omega', \mathcal{A}')$ un estadístico.

T sufficient $\iff \exists$ ^{MZ is negative} $g_0: (\mathcal{R}', \mathcal{R}') \rightarrow \mathcal{R}, \forall \theta, h: (\mathcal{R}, \mathcal{R}) \rightarrow \mathcal{R} : \varphi_0(\omega) = g_0(T\omega) \cdot h(\omega)$
 $\forall \omega, \forall \theta$

• Si B es finita, y $B \subset B' \Rightarrow B'$ finita.

• Si B e sfinite, $I \circ S = IS \Rightarrow IS$ e finte.

• T sfinite, $T = f \circ S \Rightarrow S$ sfinite.

$(\alpha, \lambda) \xrightarrow{S} (\alpha', \lambda') \xrightarrow{f} (\alpha'', \lambda'')$

$\xrightarrow{T}$

(si T suf. se poate descrie ca funcție nedele de stoc stărilor $S \Rightarrow S \text{ suf}$).

Para el caso $(\Omega, \kappa) \xrightarrow{(S_1, S_2)} \mathbb{R}^2 \xrightarrow{(P_1, P_2)} \mathbb{R}^2$, p.ej., debe ser

$$\left. \begin{aligned} T_1 &= f(S_1, S_2) \\ T_2 &= g(S_1, S_2) \end{aligned} \right\}$$

- $\mathcal{E}_n(\mathbb{R}^n, \mathcal{B}^n, \{N(\mu, \sigma^2) : \mu \in \mathbb{R}^n, \sigma^2 > 0\})$, el estadístico $(\bar{X}, S^2)$ es suficiente.

Definición: $(\Omega, \mathcal{A}, \mathcal{P})$ u.e.c.e.

a) Un estadístico real T se dice centrado si: $E_P[T] = 0 \quad \forall P \in \mathcal{P}$.

b) Una σ -álgebra $\mathcal{B}$ (o m.s. estrictamente, h.e.e., $(\Omega, \mathcal{B}, \mathcal{P})$) se dice completa si todo estadístico real centrado $\mathcal{B}$ -medible es 0 $\mathcal{P}$ -c.s.

c) Un estadístico T se dice completo si: $T^{-1}(\mathcal{A}') \ni \text{const.}$

d) Una σ -álgebra $\mathcal{B}$ se dice acotadamente completa si todo estadístico real centrado $\mathcal{B}$ -medible y acotado es 0 $\mathcal{P}$ -c.s.

• Es claro que $\mathcal{B}$ completa $\Rightarrow \mathcal{B}$ acotadamente completa.

Lema: $T: (\Omega, \mathcal{A}, \mathcal{P}) \rightarrow (\Omega', \mathcal{A}')$ completo $\Leftrightarrow (\Omega', \mathcal{A}', \mathcal{P}^T)$ completa.

• $\mathcal{B}$ completa, $\mathcal{B}' \subset \mathcal{B} \Rightarrow \mathcal{B}'$ completa. T completo, $\mathcal{S} = \mathcal{I} \circ T \equiv$ función de $T \Rightarrow \mathcal{S}$ completa.

TEORÍA DE LA DECISIÓN

Definición:

a) Un problema de decisión estadística es un par $((\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\}), (\Delta, \mathcal{D}))$ donde Θ u.e.c.e. y un espacio medible, que llamaremos espacio de decisiones.

b) Una estrategia aleatoria es una probabilidad de transición $S: \Omega \times \mathcal{D} \rightarrow [0, 1]$.

c) Una estrategia no aleatoria es una v.a. $S: (\Omega, \mathcal{A}) \rightarrow (\Delta, \mathcal{D})$.

d) Llamaremos conjunto de decisiones correctas a un subconjunto $C^* \subset \Theta \times \Delta$. Si θ es el verdadero valor del parámetro, entendemos que la decisión $S \in \Delta$ es correcta si $(\theta, S) \in C^*$.

Una estrategia no aleatoria S sugiere tomar la decisión $S(\omega) \in \Delta$ en vez de observar $\omega \in \Omega$.

Una estrategia aleatoria S sugiere tomar una decisión en Δ vía un mecanismo aleatorio regido por $S(\omega, \cdot)$, en vez de observar $\omega \in \Omega$.

Toda estrategia no aleatoria se identifica con la estr. aleatoria $S(\omega, \cdot) = \delta_{S(\omega)}$, $\delta \equiv$ delta de Dirac.

ESTIMADORES INSESGADOS DE MINIMA VARIANZA

Definición: Sea $(\Omega, \mathcal{A}, \mathbb{P}_\theta: \theta \in \Theta)$ m.e.e. y $f: \Theta \rightarrow (\mathbb{R}^k, \mathcal{B}^k)$ un estimando k -dim. Un estimador $T: (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^k, \mathcal{B}^k)$ se dice insegado si cumple que es sesgado y $E_\theta[T] = f(\theta) \forall \theta$.

Decimos que T es sesgado si no es insegado, y en ese caso la aplicamos $s: \Theta \rightarrow \mathbb{R}^k$,

$s(\theta) := E_\theta(T) - f(\theta)$ es la bias sesgo del estimador T .

• En $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P}_\theta: \theta \in \Theta)$ $\bar{X}$ y S^2 son estimadores insegados de $\mu(\theta)$ y $\sigma^2(\theta)$

(donde $\mu(\theta)$ denota la media común de la muestra de la que $(\mathbb{R}^n, \mathcal{B}^n, \mathbb{P}_\theta)$ es la e.e. conjunta).

• Hay e.e. y estimados que no admiten estimados insegados.

Definición: Una función $f: \Theta \rightarrow \mathbb{R}^k$ se dice U-estimable si existe un estimador insegado de f .

• Que T sea insegado por f implica que "la estimación que T ofrece es correcta en media", por $E_\theta[T - f(\theta)] = 0$.

(Planteamiento del problema). Sea $f: \Theta \rightarrow \mathbb{R}^k$, ¿hay algún estimador insegado de f que minimice el riesgo

definido por una cierta función de pérdida W ?

[K=1] Se suele tomar $W(\theta, x) := (x - f(\theta))^2 = \text{error cuadrático}$, cuyo riesgo es

$$\text{error cuadrático} \underset{\text{medio}}{=} E E_{M_\theta}(T) = R_T(\theta) = E_\theta[(T - f(\theta))^2] = \text{Var}_\theta(T)$$

¡El estimador óptimo sería el estimador insegado con menor Var!

[K>1] Se usa la función $W = \{W_j: j \in \mathbb{R}^k\}$, donde $W_j(\theta, x) = \left[\sum_{i=1}^n y_i (x_i - f_i(\theta)) \right]^2$.

Añ, si T y S son estimados insegados de f , $T \leq S$ si $\text{Var}_\theta(\sum y_i T_i) \leq \text{Var}_\theta(\sum y_i S_i) \forall \theta \in \Theta$. $\forall y \in \mathbb{R}^k$

Definición: Sea $f: \Theta \rightarrow \mathbb{R}^k$ una función U-estimable. Se dice que un estimador T es un estimador de mínima varianza si es insegado y prefiere a cualquier otro estimador insegado respecto al problema (*).

o Puede haber un estimador sesgado que sea preferible al EIMV de f .

Teorema (Rao - Blackwell) : Sean $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$ un e.e. y $f: \Theta \rightarrow \mathbb{R}^k$ una aplicación, y $\mathcal{B}$ una sub- σ -álgebra suficiente, X un estimador de f sumable y W una función de pérdida (en $\Theta \times \mathbb{R}^k$) continua y convexa en $\mathbb{R}^k$ para cada $\theta \in \Theta$ fijo. Entonces $E[X|\mathcal{B}]$ es un estimador de f preferible a X para el preorden def. por W . Además, si X es insesgado, $E(X|\mathcal{B})$ también lo es.

Corolario (Rao - Blackwell) : Sean $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$ un e.e. y $f: \Theta \rightarrow \mathbb{R}^k$ una aplicación, y $T: (\Omega, \mathcal{A}) \rightarrow (\Omega_1, \mathcal{A}_1)$ un estadístico suficiente, X un estimador sumable de f y W la familia de funciones de pérdida en $\Theta \times \mathbb{R}^k$ que para cada $\theta \in \Theta$ fijo son continuas y convexas en $\mathbb{R}^k$. Entonces $E(X|T) \circ T$ es un estimador de f preferible a X para el preorden def. por W .

Teorema (Lehmann - Scheffé) : Sea $\mathcal{B}$ una sub- σ -álgebra de $\mathcal{A}$ suficiente y completa, y supongamos que f admite un estimador insesgado X de miembros integrables. Entonces $E(X|\mathcal{B})$ es un estimador insesgado de f además un estimador insesgado X de miembros integrables. Entonces $E(X|\mathcal{B})$ es un estimador insesgado de f de mínima varianza de f . Además, ese estimador es esencialmente único, en el sentido de que cualquier otro estimador de mínima varianza de f es equivalente a él.

Teorema (Lehmann - Scheffé, versión para estadísticos) : Sea $T: (\Omega, \mathcal{A}) \rightarrow (\Omega_1, \mathcal{A}_1)$ un estadístico suficiente y completo para el e.e. $(\Omega, \mathcal{A}, \{P_\theta : \theta \in \Theta\})$, y X un estimador insesgado de f de miembros sumables. Entonces $E(X|T) \circ T$ es un estimador insesgado de mínima varianza de f . Además, es esencialmente único, en el sentido de que cualquier otro EIMV de f es equivalente a él.

Mejor aún, si X es función medible de T , $X = h \circ T \Rightarrow X$ es el EIMV de f .

Definición : Una estructura exponencial es un e.e. $(\mathbb{R}^k, \mathcal{B}^k, \{P_\theta : \theta \in \Theta\})$ en la que los P_θ admiten respecto a esta medida μ de $(\mathbb{R}^k, \mathcal{B}^k)$ densidades P_θ de la forma

$$P_\theta(x) = C(\theta) e^{\sum_{j=1}^s Q_j(\theta) \cdot T_j(x)} \cdot h(x), \quad x \in \mathbb{R}^k, \quad \theta \in \Theta.$$

donde $C, Q_j: \Theta \rightarrow \mathbb{R}$; $h, T_j: \mathbb{R}^k \rightarrow \mathbb{R}$, con $C, h \geq 0$. Más precisamente se dirá que es una estructura exponencial s-paramétrica.

Teorema: Sea $(\mathbb{R}^k, \mathcal{B}^k, \{P_\theta: \theta \in \Theta\})$ una estructura exponencial y $n \in \mathbb{N}$. El estadístico s -dimensional $T = (T_1, \dots, T_s)$ es suficiente para esa estructura.

• Dado un e.e. $(\mathbb{R}^k, \mathcal{B}^k, \{P_\theta: \theta \in \Theta\})$, donde $p_\theta(x) = C(\theta) e^{\sum_{j=1}^s Q_j(\theta) T_j(x)} \cdot h(x)$, cambiando la medida μ por $\nu: h = \frac{d\nu}{d\mu}$; y cambiando $Q_j(\theta)$ por una parametrización en donde $\theta'_i = Q_i(\theta)$, con $\theta'_i \in \Theta^i \subseteq \mathbb{R}^s$, entonces se puede escribir la densidad (resp. la nueva medida ν) como $p_{\theta'}(x) = C'(\theta') e^{\sum_{j=1}^s \theta'_j T_j(x)}$, reescribiendo la e.e.p. anterior como $(\mathbb{R}^k, \mathcal{B}^k, \{P_{\theta'}: \theta' \in \Theta'\})$.

Se e.e. recibe el nombre de representación natural de una estructura exponencial.

Teorema: Considera un e.exp. $(\mathbb{R}^k, \mathcal{B}^k, \{P_\theta: \theta \in \Theta\})$ s -paramétrica ($\Theta \in \mathcal{B}^s, \Theta \subseteq \mathbb{R}^s$) en su representación natural, $p_\theta(x) = C(\theta) e^{\sum_{j=1}^s Q_j(\theta) T_j(x)}$, resp. un modelo α -fijo μ .

$$\overset{\circ}{\Theta} \neq \emptyset \Rightarrow T = (T_1, \dots, T_s) \text{ es completo.}$$

En el caso en el que la representación sea la usual, $p_\theta(x) = C(\theta) e^{\sum_{j=1}^s Q_j(\theta) T_j(x)} \cdot h(x)$, ent.

$$\overset{\circ}{\left\{ (Q_1(\theta), \dots, Q_s(\theta)) : \theta \in \Theta \right\}} \neq \emptyset \Rightarrow T = (T_1, \dots, T_s) \text{ es completo}$$

$\overset{\circ}{\lim} Q$

INFERENCIA BAYESIANA

Definición: Sea $(\mathcal{A}, \mathcal{P}_\theta: \theta \in \Theta)$ un e.e., y supongamos Θ provisto de una σ -álgebra $\mathcal{T}$. Se llama distribución a priori a cualquier probabilidad Q sobre el espacio medible $(\Theta, \mathcal{T})$. A la terna $(\mathcal{A}, \mathcal{P}_\theta: \theta \in \Theta, Q)$

se le llama estructura estadística bayesiana.

• Se suele suponer que $P: \Theta \times \mathcal{A} \rightarrow [0, 1]$ es una prob. de transición, i.e., ve $P(\cdot, A): (\Theta, \mathcal{T}) \rightarrow [0, 1]$ es medible. En particular, esto ocurre cuando $L(\omega, \theta)$ sea $\mathcal{A} \times \mathcal{T}$ -medible (v.a.r.).

Teoreme: Soit un e.e. bayésien $(\Omega, \mathcal{A}, \{P_\theta: \theta \in (\Theta, \mathcal{T}, \mathcal{Q})\})$. Soit donc

$$\beta_\theta^*: \mathcal{A} \longrightarrow [0, 1]$$

$$A \longmapsto \beta_\theta^*(A) := \int P_\theta(A) dQ \quad \dots \quad (\text{se intègre resp. } \theta \text{ en } \Theta).$$

une probabilité qu'on appelle probabilité prédictive. Si aussi, $\{P_\theta\} \ll \mu, \Rightarrow \beta_\theta^* \ll \mu$, et si la fonction de vraisemblance est v.a. (i.e., $\mathcal{A} \times \mathcal{T}$ -mesurable), alors $\frac{d\beta_\theta^*}{d\mu}(\omega) = \int P_\theta(\omega) dQ$

Définition: Supposons qu'il existe une probabilité de transition $P^*: \Omega \times \mathcal{T} \longrightarrow [0, 1]$ tel que

$$\int_{\mathcal{T}} P_\theta(A) dQ = \int_{\mathcal{A}} P_\omega^*(T) d\beta_\theta^* \quad \text{Enters pour each observation } \omega, \text{ se appelle } \underline{\text{distribution a posteriori}} \text{ ou}$$

la probabilité P_ω^* sobre $(\Theta, \mathcal{T})$.

" $P^*(\omega, \cdot)$ "

Teoreme: Soit $(\Omega, \mathcal{A}, \{P_\theta: \theta \in (\Theta, \mathcal{T}, \mathcal{Q})\})$ un e.e.-s. donné par une mesure σ -finie μ . Soit $P_\theta = \frac{dP_\theta}{d\mu}$ et sup. que $L(\omega, \theta)$ est une v.a. Alors existent les distributions a posteriori P_ω^* , et aussi:

$$\frac{dP_\omega^*}{dQ}(\theta) = \frac{f_\theta(\omega)}{\int f_\theta(\omega) dQ(\theta')}, \quad \beta_\theta^* \text{-c.s.}$$

Aussi, si Π est la prob. sobre $\mathcal{A} \times \mathcal{T}$: $\Pi(A \times T) = \int_{\mathcal{T}} P_\theta(A) dQ$ (une extension naturelle de la Mes. Prod. Généralisée), et $\mathbf{I}$ et $\mathbf{J}$ sont les projecteurs de $\Omega \times \mathcal{T}$ en Ω et $\mathcal{T}$, resp; alors

$$P_\omega^* = \Pi^{\mathbf{J}} \mathbf{I} = \omega$$

$$(\Omega, \mathcal{A}, \{P_\theta: \theta \in (\Theta, \mathcal{T}, \mathcal{Q})\}) \equiv (\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi).$$

Définition: Un problème de décision bayésien est une triple

$$\left((\Omega, \mathcal{A}, \{P_\theta: \theta \in (\Theta, \mathcal{T}, \mathcal{Q})\}), (\Delta, \mathcal{D}), \leq \right)$$

$\uparrow$ $\uparrow$ $\uparrow$
 e.e.-s. $\uparrow$ $\uparrow$ $\uparrow$
 exp. de dec. $\uparrow$ $\uparrow$ $\uparrow$
 prior

Defin:

a) Una función de pérdida es una v.a. $W: (\Theta \times \Delta, \mathcal{T} \times \mathcal{Q}) \rightarrow [0, \infty]$.

b) Sea S una estrategia y W función de pérdida. Se llama riesgo medio o riesgo de Bayes a

$$R_S^Q := \int R_S(\theta) dQ = \int_{\Theta} \int_{\Delta} W(\theta, S) S(\omega, dS) dP_{\theta} dQ.$$

Defin: Dada una función de pérdida W , se llama estrategia de Bayes a toda estrategia S que minimice R_S^Q .

• Una función de pérdida W define un preorden entre las estrategias:

$$S \preceq_W S' \iff R_S^Q \leq R_{S'}^Q$$

• Dos estrategias que difieren en componentes idénticos:

• En Estadística Bayesiana, el problema de estimar $f: \Theta \rightarrow \mathbb{R}$ reduce a $T: (\Omega, \mathcal{X}) \rightarrow \mathbb{R}$ dada una función de pérdida W constante en cuanto a una estrategia que minimice el riesgo Bayes.

Teorema: Sea $(\Omega, \mathcal{X}, \mathcal{P}_0: \Theta \subset (\Theta, \mathcal{T}, \mathcal{Q}))$ es dado por μ_i y sea $\mathcal{P}_0 = \frac{dP_{\theta}}{d\mu}$. Sup. que $\mathcal{L}$ es una v.a. y sea Π como antes. Sea $f: (\Theta, \mathcal{T}) \rightarrow \mathbb{R}$ una v.a. con varianza finita a priori, y sea la aplicación

$$F: \Omega \times \Theta \rightarrow \mathbb{R}, F(\omega, \theta) = f(\theta), \text{ y } W(\theta, x) = (f(\theta) - x)^2.$$

Entonces, la esperanza condicional $E_{\Pi}[F|I] \stackrel{\text{int}}{=} E_{\Pi}[f(\theta)|\omega] \approx$ el estimador Bayes de f .

delong $\Rightarrow$

$$E_{\Pi}[f(\theta)|\omega] = \frac{\int f(\theta) \mathcal{P}_0(\omega) dQ}{\int \mathcal{P}_0(\omega) dQ} = E_{P_{\omega}^*}[f].$$

y es óptimo P_{ω}^* -c.s.

INTERVALOS DE CONFIANZA

• Si queremos estimar un cierto parámetro, ya sabemos cómo hacerlo (EMV), pero ahora queremos también dar un intervalo que dé una idea del error.

Definición: Sea $(\Omega, \mathcal{A}, P_0, \theta \in \Theta)$ m.e.e. y $f: \Theta \rightarrow \Theta'$ una aplicación.

a) Un conjunto de confianza para f es una aplicación $S: \Omega \rightarrow \mathcal{P}(\Theta') : \{\omega: f(\theta) \in S(\omega)\} \in \mathcal{A} \forall \theta$.

b) Si $0 < \alpha < 1$, un conjunto de confianza al nivel (de confianza) $1-\alpha$ es un conj. de conf. S :

$$P_0(\{\omega: f(\theta) \in S(\omega)\}) \geq 1-\alpha \quad \forall \theta \in \Theta.$$

• Lo anterior significa que con P_0 -probabilidad $\geq 1-\alpha$, el conjunto $S(\omega)$ contiene a $f(\theta)$; y es $\forall \theta \in \Theta$ (en particular para el verdadero valor del parámetro).

¡IMPORTANTE: Sea $S(\omega) = [\omega-2, \omega+2] \subset \mathbb{R}$, un conj. de conf. al nivel $0,95 = 1-\alpha$

- NO es correcto decir que, si he visto $\omega = 50$ (valor observado), que " θ está entre 48 y 52 con prob. superior al 95%". ¡Esto asegura que Θ tiene una probabilidad!

- SI es correcto decir que $[\omega-2, \omega+2]$ contiene al parámetro θ con P_0 -prob. $\geq 0,95$, y es $\forall \theta \in \Theta$.

- (*) SI es correcto (versión frecuentista) decir que, al repetir el experimento un número grande de veces, el 95% de los valores $\omega \in \mathbb{R}$ observados dan lugar a intervalos $[\omega-2, \omega+2]$ que contienen al verdadero valor del parámetro.

* Esto se lo interpret. como de la afirmación $P_0(\{\omega: \theta \in [\omega-2, \omega+2]\}) \geq 0,95$.

¿Cómo construir conjuntos de confianza?

Definición: Sea $(\Omega, \mathcal{A}, P_0)$ m.e.e. y $f: \Theta \rightarrow \mathbb{R}$ una función a estimar. Se dice que una aplicación $T: \Omega \times \Theta \rightarrow \mathbb{R}$ es una cantidad pivote para f si

i) $\forall \theta \in \Theta, T(\cdot, \theta): \Omega \rightarrow \mathbb{R}$ es v.a.r.

ii) $P_0^{T(\cdot, \theta)}$ no depende de $\theta \in \Theta$.

Notar que si $\alpha \in (0,1)$, $\exists a_1, a_2: P_0^{T(\cdot, \theta)}[a_1, a_2] \geq 1-\alpha \forall \theta$.

SI es correcto decir que, para cada $\theta \in \Theta$, $[a_1, a_2]$ es un intervalo de P_0 -prob. $\geq 1-\alpha$; o que el valor observado ω está en $[a_1, a_2]$ con una P_0 -prob. $\geq 1-\alpha$.

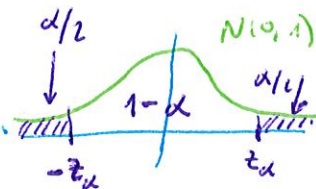
Proposición: Sea T una cantidad pivote para f , y supongamos que para cada $\omega \in \Omega$ hay números $\phi_1(\omega)$, $\phi_2(\omega)$: para cada θ ,

$$\alpha_1 \leq T(\omega, \theta) \leq \alpha_2 \iff \phi_1(\omega) \leq f(\theta) \leq \phi_2(\omega).$$

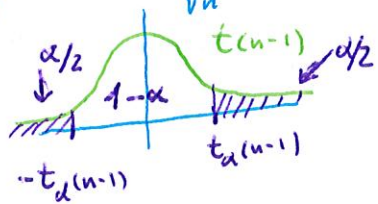
Entonces $S(\omega) = [\phi_1(\omega), \phi_2(\omega)]$ es un intervalo de confianza al nivel $1-\alpha$ de f .

Principio de sustitución: A los parámetros desconocidos, sustituir uno de ellos por un estimador de él (ej.: σ^2 por S^2).

Ej.: $\bar{X} \pm z_\alpha \frac{\sigma_0}{\sqrt{n}}$ es un int. de conf. para μ al nivel $1-\alpha$ (modelo normal, σ_0 conocido).



Ej.: $\bar{X} \pm t_{\alpha(n-1)} \frac{S}{\sqrt{n}}$ es un int. de conf. de μ al nivel $1-\alpha$ (μ y σ desconocidos).



III: CONTRASTE DE HIPÓTESIS

• Un contraste de hipótesis consiste en tomar la decisión de aceptar o rechazar una cierta afirmación sobre la distribución desconocida de una v.a. X a partir de observaciones. Se fija $P = \{P_0\}$ que divide entre los que tienen las posibles distribuciones de X que hacen cierta la afirmación y los que la hacen falsa. Tal afirmación puede entonces identificarse por un subconjunto de P .

TEST DE HIPÓTESIS

Definición: Sea $(\Omega, \mathcal{A}, \{P_0, P_1\} \subset \Theta)$. Una hipótesis es un subconjunto propio de Θ . Si es unitario se dice simple, si no compuesta. Se llama hipótesis nula a la hipótesis a contrastar; y la denotamos por Θ_0 . El complemento $\Theta_1 := \Theta_0^c$ es la hipótesis alternativa.

Definición: Un test es un estadístico $\phi: (\Omega, \mathcal{A}) \rightarrow [0, 1]$. Si $\phi(\omega) = 0$ o $1 \forall \omega$ (i.e. si $\phi = I_C, C \in \mathcal{A}$) se dice que es no aleatorio. En tal caso, se llama región crítica a $C = \{\omega: \phi(\omega) = 1\}$ y región de aceptación a $C^c = \{\omega: \phi(\omega) = 0\}$.

Definición: Sea $(\Omega, \mathcal{A}, \{P_0, P_1\} \subset \Theta)$ una cc y Θ_0 una hipótesis. El problema de contrastar la hipótesis nula Θ_0 contra la alternativa es el problema de decisión estadística definido por su estructura estadística y el espacio de decisiones $(\Delta, \mathcal{D})$, con $\Delta = \{\Theta_0 \stackrel{\text{int}}{=} 0, \Theta_1 \stackrel{\text{int}}{=} 1\}$, $\mathcal{D} = \mathcal{P}(\Delta)$.

Definición: Si ϕ es un test de Θ_0 vs Θ_1 , se llama función potencia del test ϕ a

$$\begin{aligned} \beta_\phi: \Theta &\rightarrow \mathbb{R} \\ \theta &\mapsto \beta_\phi(\theta) := E_\theta[\phi]. \end{aligned}$$

: $\theta_1 \in \Theta_1$, a $\beta_\phi(\theta_1)$ se le llama potencia de ϕ contra la alternativa Θ_1 .

Definición: Dos test ϕ, ψ ^{de Θ_0} ~~centro Θ_1~~ se dicen equivalentes si tienen la misma función potencia.

Una estrategia aleatoria para un problema de contraste de hipótesis queda determinada por un test ϕ : el test ϕ sugiere se, observando $\omega \in \Omega$, aceptemos la hipótesis nula con prob. $1 - \phi(\omega)$ y rechacemos con $\phi(\omega)$, i.e.,

$$S(\omega, \{0, 1\}) = 1 - \phi(\omega) \quad ; \quad S(\omega, \{1, 0\}) = \phi(\omega).$$

Entonces, toda estrategia S es la estrat. def por un test: basta definir como $\phi(\omega) := S(\omega, \{1, 0\})$.

• Un conjunto de decisiones correctas $\Rightarrow C^* = \{(\theta_0, 0) : \theta_0 \in \Theta_0\} \cup \{(\theta_1, 1) : \theta_1 \in \Theta_1\}$.

• (Errors de tip I y II)

	Θ_0 cierto	Θ_0 falso
Aceptar Θ_0	Decisión correcta	Error tipo II
Rechazar Θ_0	Error tipo I	Decisión correcta

¡¡¡ θ_0 es un valor que hace cierto la hipótesis!!!!

• Si ϕ es test u decisión, para cada $\theta \in \Theta$,

$$\text{prob. } P_0 \text{ rechazar } \Theta_0 = \text{prob. } P_1 \text{ aceptar } \Theta_1 = P_0(\phi=1) = \beta_\phi(\theta),$$

luego si $\theta_0 \in \Theta_0$, $\beta_\phi(\theta_0)$ representa una probabilidad de error de tipo I, $\theta_1 \in \Theta_1$, $1 - \beta_\phi(\theta) = 1 - P_0(\phi=1) = P_0(\phi=0) = P_0(C)$ representa una prob. de error de tipo II.

• El valor estadístico de un test queda determinado por su función potencia.

• En general no es posible minimizar los errores de tipo I y II simultáneamente. Es preferible controlar los errores de tipo I, pues la función de la hipótesis se basa en estudios anteriores. Para demostrar los estudios nuevos se debe ser muy seguros! Más que conjeturas de razón!

Proposición: Sea ϕ un test Θ_0 vs Θ_1 , y B sub- σ -álgebra. Entonces $E(\phi|B)$ es un test equivalente a ϕ .

Lema: Sea ϕ un test Θ_0 vs Θ_1 , y $T: (\mathcal{X}, \mathcal{A}) \rightarrow (\mathcal{Y}, \mathcal{B})$ un estadístico inf. para $\{P_\theta : \theta \in \Theta\}$. Entonces existe una v.a. $\varphi: (\mathcal{Y}, \mathcal{B}) \rightarrow [0, 1]$: $\varphi \circ T$ es un test equivalente a ϕ .

Definición: Sea $(\mathcal{X}, \mathcal{A}, \{P_\theta\}) \ll \mu$ y $P_0 = \frac{dP_0}{d\mu}$. Se dice que ϕ es un test de la razón de verosimilitudes para contrastar Θ_0 vs Θ_1 si $\exists C > 0$:

$$\lambda(w) < C \Rightarrow \phi(w) = 1$$

$$\lambda(w) > C \Rightarrow \phi(w) = 0$$

$$\text{donde } \lambda(w) = \frac{\sup_{\theta_1 \in \Theta_1} L(w, \theta_1)}{\sup_{\theta_0 \in \Theta_0} L(w, \theta_0)}$$

• El C se puede calcular fijando $\alpha \in (0, 1)$ y exigiendo el test verifique $\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\{w : \lambda(w) < C\}) = \alpha$.

Si no se puede así, se hace $\phi(w) = \gamma$ si $\lambda(w) = C$ y se determinan C y γ :

$$\sup_{\theta_0 \in \Theta_0} \left(P_{\theta_0}(\{w : \lambda(w) < C\}) + \gamma P_{\theta_0}(\{w : \lambda(w) = C\}) \right) = \alpha.$$

Definición: Sea ϕ un test Θ_0 vs Θ_1 . Se llama extensión de ϕ a $\alpha_\phi := \sup_{\theta_0 \in \Theta_0} \beta_\phi(\theta_0) \in (0, 1)$.

Definición: Sea $\alpha \in (0, 1)$ y Φ_α sea la clase de los test ϕ Θ_0 vs Θ_1 de extensión $\leq \alpha$.

a) Si $\phi, \phi' \in \Phi_\alpha$, se dice que ϕ es uniformemente más potente que ϕ' si $\beta_\phi(\theta_1) \geq \beta_{\phi'}(\theta_1) \forall \theta_1 \in \Theta_1$.

b) Si existe un test $\phi \in \Phi_\alpha$ más potente que cualquier otro test en Φ_α , se dice que ϕ es un test UMP(α) de Θ_0 vs Θ_1 (test UMP al nivel α).

Lema (Fundamental de Neyman-Pearson): Sea $\alpha \in (0, 1)$ y P_0 y P_1 prob. en $(\Omega, \mathcal{A})$, con densidad f_0 y f_1 , resp. una medida dominante μ . ($\Theta = \{0, 1\}$).

1) (Existencia): Para contrastar la hipótesis P_0 vs P_1 existe un test ϕ y una $K > 0$:

$$(NP1) \quad \beta_\phi(0) = E_0(\phi) = \alpha$$

$$(NP2) \quad \phi(\omega) = \begin{cases} 1 & \text{si } f_1(\omega) > K f_0(\omega) \\ 0 & \text{si } < \end{cases}$$

2) (Condición suficiente para que un test P_0 vs P_1 sea UMP): Si ϕ satisface (NP1) + (NP2) para algún $K \in \mathbb{R} \Rightarrow \phi$ es UMP(α) de P_0 vs P_1 .

3) (Condición necesaria para que un test P_0 vs P_1 sea UMP): Si ϕ es UMP(α) de P_0 vs P_1 , entonces existe $K > 0$: (NP2) se cumple p.c.s. Además, ϕ cumple (NP1) a no ser que exista un test de extensión $< \alpha$ y potencia 1 contra P_1 .

Corolario: Sea ϕ es test UMP(α), $\alpha \in (0, 1)$, de P_0 vs P_1 . Si β es la potencia de ϕ contra P_1 , entonces $\alpha < \beta$ (a menos que $P_0 = P_1$).

$$\begin{pmatrix} \beta = \beta_1(\phi) \\ \text{"} \\ E_1(\phi) \end{pmatrix}$$

Definición: Sea $(\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$ me ee domado por μ_0 -finit, y sea $P_0 = \frac{dP_0}{d\mu}$. Se dice que tal ee tiene razón de verosimilitud monotona si

i) $\Theta \subseteq \mathbb{R}$

ii) Existe un estadístico $T: (\Omega, \mathcal{A}) \rightarrow \mathbb{R}$: si $\theta < \theta' \Rightarrow P_\theta \neq P_{\theta'}$ y $\exists g_{\theta, \theta'}: \mathbb{R} \rightarrow \mathbb{R}$

creciente : $P_{\theta'}(w) = g_{\theta, \theta'}(T(w)) \cdot P_\theta(w)$, $\forall w \in \Omega$.
(w dec.)

Se dice también que h ee tiene razón de verosimilitud monotona creciente en el estadístico real T . (RVM)

Teorema (Karlin - Lehmann - Rubin): Sea $(\Omega, \mathcal{A}, \{P_\theta\})$ me ee con R.V.M creciente en el estadístico real T , y $\Theta_0 \in \Theta$.

1) Dado $\alpha \in (0, 1)$ existe un test ϕ UMP(α) para $\theta \leq \theta_0$ vs. $\theta > \theta_0$, definido por

(RVM1)
$$\phi(w) = \begin{cases} 1, & T(w) > C \\ \gamma, & T(w) = C \\ 0, & T(w) < C \end{cases}$$

donde C, γ son ites determinados por

(RVM2)
$$E_{\theta_0}(\phi) = \alpha$$

2) La función potencia $\beta_\phi(\theta)$ del test e estrictamente creciente en $\{\theta: 0 < \beta_\phi(\theta) < 1\}$.

3) Para cada $\theta' \in \Theta$, el test de 1) e UMP al nivel $\alpha' = \beta_\phi(\theta')$. (para contrastar $\theta \leq \theta'$ vs. $\theta > \theta'$)

• También podemos hablar de RVM decreciente, y en el Th. hecho se cambiar $<$ por $>$. Si es creciente, si quisieramos comparar $\theta \geq \theta_0$ vs. $\theta < \theta_0$, también debemos cambiar $>$ por $<$ y $<$ por $>$ en el Th. de 1).

$$\phi(w) = \begin{cases} 1, & T(w) > C \\ \gamma, & T(w) = C \\ 0, & T(w) < C \end{cases}$$

$$; \quad \phi(w) = \begin{cases} 1, & T(w) < C \\ \gamma, & T(w) = C \\ 0, & T(w) > C \end{cases}$$

Cre $\theta \leq \theta_0$ vs $\theta > \theta_0$

o Dec $\theta \geq \theta_0$ vs $\theta < \theta_0$

Dec $\theta \leq \theta_0$ vs $\theta > \theta_0$

o Cre $\theta \geq \theta_0$ vs $\theta < \theta_0$

TEST INVARIANTES

Supongamos que $(\Omega, \mathcal{A}, \mathbb{P}_0)$ es identificable, es decir, $P_0 \neq P_0'$ si $\theta \neq \theta'$.

Definición: Una transformación en Ω es una biyección $\Omega \rightarrow \Omega$.

Definición: Sea $\mathcal{G}$ un grupo de transformaciones en Ω . Se dice que $(\Omega, \mathcal{A}, \mathbb{P}_0)$ es invariante bajo la acción del grupo $\mathcal{G}$ si:

- i) $\forall g \in \mathcal{G}$ es biyectiva
- ii) $\forall g \in \mathcal{G}, \forall \theta \in \Theta, P_\theta^g \in \mathcal{P}_0 \mathcal{G}$.

Denotamos con $\bar{g}(\theta)$ al único punto: $P_\theta^g = P_{\bar{g}(\theta)}$.

Proposición: Sea $\mathcal{G}$ un grupo de trnsf que dejan invariante $(\Omega, \mathcal{A}, \mathbb{P}_0)$. Si $g_1, g_2 \in \mathcal{G}$,

- 1) $\bar{g}(\emptyset) = \emptyset$
- 2) $\bar{g}$ es una biyección $\emptyset \rightarrow \emptyset$
- 3) $\overline{g_1 \circ g_2} = \bar{g}_1 \circ \bar{g}_2$
- 4) $\overline{g_1^{-1}} = \bar{g}_1^{-1}$

El grupo $\bar{\mathcal{G}} := \{\bar{g} : g \in \mathcal{G}\}$ de transformaciones de $\emptyset$ se llaman grupos asociados.

Definición: Sea $(\Omega, \mathcal{A}, \mathbb{P}_0 : \theta \in \Theta)$ un ee y $\emptyset_0$ un hipotesis en $\emptyset$. Se dice que el grupo $\mathcal{G}$ deja invariante el problema de contrastar $\emptyset_0$ vs $\emptyset_1$ en su estructura si $\mathcal{G}$ deja invariante la estructura si

$$\forall g \in \mathcal{G} \quad \bar{g}(\emptyset_0) = \emptyset_0, \quad \bar{g}(\emptyset_1) = \emptyset_1.$$

Définition: Soit $(\Omega, \mathcal{A}, P_0, P_1)$ une ee $\mathcal{G}$ un gpp qui laisse invariants Θ_0 vs Θ_1 . Se dit que ϕ est un test invariant par $\mathcal{G}$ si pour cada $\omega \in \Omega$ y $g \in \mathcal{G}$, $\phi(g(\omega)) = \phi(\omega)$.

Définition: Un test ϕ pour el problema $\mathcal{G}$ -invariant Θ_0 vs Θ_1 en $(\Omega, \mathcal{A}, P_0, P_1)$ se dice UMP invariant al nivel α ($UMPI(\alpha)$) si ϕ es un test invariant de extension $\leq \alpha$ y ^{uniforme} ~~stricto~~ ^{potente} que cualquier otro test de extension $\leq \alpha$.

Définition: Une application $M: \Omega \rightarrow \Omega$ se dice invariante par $\mathcal{G}$ si $M(g(\omega)) = M(\omega) \quad \forall g \in \mathcal{G}, \forall \omega \in \Omega$.
Se dice que M es invariante maximal si es invariante y $\mathcal{G}$ -adaptada: $M(\omega_1) = M(\omega_2) \Rightarrow \exists g \in \mathcal{G}: g(\omega_1) = \omega_2$.

Teorema: Soit M une application invariante maximal app. au gpp $\mathcal{G}$, y sea ϕ un test.

$$\phi \text{ invariante} \iff \exists h: \Omega \rightarrow [0, 1]: \phi(\omega) = h(M(\omega)) \quad \forall \omega \in \Omega.$$

$$\phi = h \circ M$$

Teorema: Soit $(\Omega, \mathcal{A}, P_0, P_1)$ ee $\mathcal{G}$ gpp de transf. invariants sobre elle, que le deje invariante, y sea $M: (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{A}')$ un estadístico invariante por $\mathcal{G}$. Si $v: \Theta \rightarrow \Theta'$ es una aplic. invariante maximal por $\overline{\mathcal{G}}$, entonces P_0^M solo depende de Θ via $v(\theta)$.

MODELO LINEAL NORMAL

• Se pretende estudiar los efectos producidos por un unico factor que actúa sobre un cierto número de niveles (ej: pérdida de peso producidos por varios dietas)

Définition: Un problema de inferencia estadística se dice lineal normal si se ee $\rightarrow$

$$(\mathbb{R}^n, \mathcal{B}^n, \{N(\mu_1, \sigma^2) \times \dots \times N(\mu_n, \sigma^2) : (\mu_1, \dots, \mu_n) \in V, \sigma^2 > 0\}) \quad (*)$$

donde V es un subespacio vectorial de $\mathbb{R}^n$ de dim $s < n$.

Définition: Soit $1 \leq r \leq s < n$, y V sev de dim s , y W un sev de V de dim $s-r$. Un problema lineal de test de hipótesis consiste en contrastar la hipótesis $(\mu_1, \dots, \mu_n) \in W$ en la ee anterior (**)

• Tales pruebas ortogonales pueden calcularse con métodos estadísticos a través de la notación

Linea: Supongamos que tenemos datos en un modelo de clas. simple de A.V.

$$\underbrace{x_{11}, \dots, x_{1n_1}, \dots, x_{s1}, \dots, x_{sn_s}}_{\substack{x_{i0} \quad x_{s0} \\ \leftarrow \text{medias en cada "tratamiento"} \\ \leftarrow \text{media total}}}$$

Entonces $\tilde{\mu}(x) = (x_{..}, \dots, x_{..})$; $\hat{\mu}(x) = (x_{10}^{n_1}, \dots, x_{10}^{n_1}, \dots, x_{s0}^{n_s}, x_{s0}^{n_s})$.

Teorema: El test UMPI(α) en un modelo de clasificación simple de análisis de la varianza

$$(R^n, B^n, \{N(\mu_i, \sigma^2)\}^{n_1} \times \dots \times N(\mu_s, \sigma^2)^{n_s} : \mu_1, \dots, \mu_s \in \mathbb{R}, \sigma^2 > 0)$$

para la hipótesis de igualdad de medias $H_0: \mu_1 = \dots = \mu_s \Rightarrow$

$$\frac{\frac{1}{s-1} \sum_{i=1}^s \sum_{j=1}^{n_i} (x_{i0} - x_{i0}^e)^2}{\frac{1}{n-s} \sum_{i=1}^s \sum_{j=1}^{n_i} (x_{ij} - x_{i0})^2} > F_\alpha(s-1, n-s).$$

• El Te. Pitégoras dice que

$$\|x - \tilde{\mu}(x)\|_2^2 = \|x - \hat{\mu}(x)\|_2^2 + \|\hat{\mu}(x) - \tilde{\mu}(x)\|_2^2$$

$$\underbrace{\sum_{i=1}^s \sum_{j=1}^{n_i} (x_{ij} - x_{..})^2}_{\text{variancia total}} = \underbrace{\sum_{i=1}^s \sum_{j=1}^{n_i} (x_{ij} - x_{i0})^2}_{\text{variancia "dentro"}} + \underbrace{\sum_{i=1}^s n_i (x_{i0} - x_{i0}^e)^2}_{\text{variancia "entre" (las distintas poblaciones)}}$$

variancia total

variancia "dentro"

variancia "entre" (las distintas poblaciones)

(variabilidad explicada por el error)
(variancia procedente del interior de las distintas poblaciones)

(variabilidad explicada por el error y los efectos de los factores)

Definici: Direm que les problemes

Contrasta Θ_0 vs Θ_1 en
 $(\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$

Contrasta Θ'_0 vs Θ'_1 en
 $(\Omega', \mathcal{A}', \{P'_\theta: \theta \in \Theta'\})$

són equivalents si existe una bijecció bijectiva $\mu: (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ i una bijectió $m: \Theta \rightarrow \Theta'$ tal que
 $P'_\theta = P'_{m(\theta)} \quad \forall \theta \in \Theta, \quad m(\Theta_0) = \Theta'_0; \quad m(\Theta_1) = \Theta'_1.$

Teorema: El problema $(*)$ per contrastar la hipòtesi $(**)$ és equivalent a $(\mathbb{R}^+, \mathcal{B}^+, \{U_\lambda: \lambda > 0\})$

amb la hipòtesi $H_0: \lambda = 0$, que llavors, problema uniforme en λ

Lema: Si $\lambda > 0$, la distribució U_λ és abs. cont. resp. Lebesgue, i més precisament $f_\lambda > 0$ en $\mathbb{R}_+$. Si $\lambda > 0$,

f_λ / f_0 és estrictament creixent.

Teorema: Sea X_i les observacions de $(*)$ i $Y = OX$, donde O és una matriu ortogonal. La regió UMPI(α)

per $H_0: \lambda = 0 \rightarrow$

$$Z^* = \frac{\frac{1}{r} \sum_{i=1}^r Y_i^2}{\frac{1}{n-s} \sum_{i=s+1}^n Y_i^2} > F_\alpha(r, n-s).$$

Corol·lari: Si $F(x_1, \dots, x_n) := Z^*(y_1, \dots, y_n)$, $y_i = \sum O_{ij} x_j$, $O = (O_{ij})$, Entons el test definit en $(*)$ per la regió crítica

$$F > F_\alpha(r, n-s)$$

és UMPI(α) per contrastar $(**)$. A més

donde $\hat{\mu}(x)$ és la proy. ortogonal de x sobre V i $\tilde{\mu}(x)$ sobre W . Per tant

$$\sum_{i=1}^r Y_i^2 = \|x - \hat{\mu}(x)\|_2^2, \quad \text{i} \quad \sum_{i=s+1}^n Y_i^2 = \|\hat{\mu}(x) - \tilde{\mu}(x)\|_2^2,$$

$$F(x) = \frac{\frac{1}{r} \|\tilde{\mu}(x) - \hat{\mu}(x)\|_2^2}{\frac{1}{n-s} \|x - \hat{\mu}(x)\|_2^2}.$$



¿Pq $x_{11} \neq x_{10}$? Taché han
residuo de una dieta, luego
var.



¿Pq $x_{11} \neq x_{55}$?
Habrá error... ii pero sobre todo
siempre q han residuo distintos dietas!!

• En ~~(X)~~,

1. $\hat{\mu}$ EIMV de μ

2. $\frac{1}{n-3} \|x - \hat{\mu}(x)\|^2$ EIMV de σ^2 .

