ESTADISTICA

1. MODELOS PROBABILÍSTICOS DISCRETOS

PROBABILIDAD EMPIRICA: cociente entre si un acontecimiento ocurre o no y el número de repeticiones del experimento observado

$$P(A) = \frac{n(A)}{n}$$

3 propiedades: $0 \leq n(A) \leq n \rightarrow 0 \leq P(A) \leq 1$

$n(A) = n \rightarrow P(A) = 1$

$n(A \cup B) = n(A) + n(B) \rightarrow P(A \cup B) = P(A) + P(B)$

EXPERIMENTO ALEATORIO: no se conoce su resultado de antemano, el conjunto de sus posibles resultados está bien determinado y se puede repetir en idénticas condiciones siempre que se quiera

↳ SUCESO: acontecimiento que ocurre o no según el resultado de la prueba ↴ subconjunto de

$A \subset \Omega$

↳ ESPACIO MUESTRAL: conjunto de resultados posibles

Ω

↓ proposiciones lógicas

- suceso imposible: $A = \emptyset$
- suceso seguro: $A = \Omega$
- sucesos incompatibles/disjuntos: $A \cap B = \emptyset$

PROBABILIDAD: $P \equiv$ probabilidad $\mathcal{A} \equiv$ álgebra sucesos $\Omega \equiv$ espacio muestral → espacio probabilístico

$P : \mathcal{A} \to [0,1]$, $A \in \mathcal{A} \to P(A)$

se cumple:
1. $P(\Omega) = 1$
2. A, B disjuntos $P(A \cup B) = P(A) + P(B)$ → aditividad

- ω, suceso elemental de Ω. $0 \leq P(\omega) \leq 1$ y $\sum_{\omega \in \Omega} P(\omega) = 1$

entonces: $P(A) = \sum_{\omega \in A} P(\omega)$

PROPIEDADES PROBABILIDAD: sucesos A, A^c disjuntos → $A \cup A^c = \Omega$

sobre $(\Omega, \mathcal{A}, P)$

→ comprobables mediante diagramas de Venn

- probabilidad suceso contrario: $\boxed{P(A^c) = 1 - P(A)}$
- probabilidad diferencia: $\boxed{P(A - B) = P(A) - P(A \cap B)}$
- probabilidad unión sucesos: $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$
- " 3 sucesos: $P(A_1 \cup A_2 \cup A_3) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + P(A_1 \cap A_2 \cap A_3)$
- propiedad monotonía: $B \subset A \to P(B) \leq P(A)$

ASIGNACIÓN PROBABILIDADES

- suceso elemental: $\frac{1}{\#(\Omega)}$; $\#(\Omega)$ → cardinal de Ω → $card(\Omega)$
- suceso general: $\boxed{P(A)} = \sum_{\omega \in A} \frac{1}{\#(\Omega)} = \frac{\#(A)}{\#(\Omega)} = \boxed{\frac{\text{nº casos favorables}}{\text{nº casos posibles}}}$

regla de Laplace

- modelo equiprobable: los sucesos elementales de Ω tienen la misma probabilidad → un dado no cargado, una moneda no cargada ...

CALCULO POR COMPLEMENTARIO: para sucesos definidos por la condición "al menos" → probabilidad suceso -1

MODELOS DINAMICOS: experimentos compuestos de varios sucesos que ocurren al mismo tiempo → cada subexperimento influye en las condiciones en que ocurren los posteriores

PROBABILIDAD CONDICIONADA: $\boxed{P(B|A) = \frac{P(A \cap B)}{P(A)}}$ → suceso B tal que "si A" ó "sabiendo que A"

↓

P(A) es conocida y condiciona el resto de sucesos

prob. posterior — prob. previa

3 propiedades

$P(A|A) = 1$

$P(B^c|A) = 1 - P(B|A)$

$P(B_1 \cup B_2|A) = P(B_1|A) + P(B_2|A)$ con B_1, B_2 disjuntos

- cálculo dinámico: $\boxed{P(A \cap B) = P(A)\,P(B|A)}$
- fórmula probabilidad total: $\boxed{P(A) = \sum_{i=1}^{n} P(B_i)\,P(A|B_i)}$ → B_i, B_j disjuntos
  ↳ fracciona el espacio muestral en B_i sucesos
- fórmula de Bayes: $\boxed{P(A_j|B) = \frac{P(A_j)\,P(B|A_j)}{\sum_{i=1}^{n} P(A_i)\,P(B|A_i)}}$ → conocemos el resultado de un subexperimento y no se sabe el de otro subexperimento anterior

SUCESOS INDEPENDIENTES: repetir un experimento sin que el resultado influya en los restantes

- suceso A favorable a B si $P(B|A) > P(B)$: B es más frecuente cuando A ocurre
- suceso A independiente de B si $P(B|A) = P(B)$: A y B son independientes
  ↳ extensible a n sucesos
  ↳ también lo son sus complementarios

$\boxed{P(A \cap B) = P(A) \cdot P(B)}$ ↳ contrario a ser disjuntos

extensible a experimentos independientes (lanzar una moneda y luego un dado) componiendo los espacios muestrales $\Omega_1 \times \Omega_2$

VARIABLES ALEATORIAS: X sobre un espacio probabilístico (Ω, P) es una función $X: \Omega \to \mathbb{R}$

↳ los sucesos elementales se asocian a la observación de la variable X

- cuando $X(\Omega)$ es finito, la variable es DISCRETA

↳ asignando un único valor numérico a cada $\omega \in \Omega$

↓

DISTRIBUCIÓN DE PROBABILIDAD : modo en que se reparte la probabilidad entre los posibles valores de la variable

$P(X=x) = p(x) \ \forall x \in X(\Omega)$

x	- - -
$p(x)$	- - -

↓

$\sum p(x) = 1$

↓

$x \to p(x)$: función de probabilidad

• si X es discreta, $Y = g(X)$, Y es discreta

VALOR ESPERADO : esperanza matemática, promedio de los valores que toma la variable, ponderado por el valor de su probabilidad ("centro de gravedad" de la distribución de probabilidad)

$$\boxed{E(x) = \sum_{x \in X(\Omega)} x \, P(X=x)}$$

MOMENTOS VARIABLE : son los valores esperados de las funciones potenciales de una variable y describen la forma de su distribución

$$\mu_r = E(X^r) \quad , r > 0$$

o los valores esperados de las potencias de las desviaciones respecto de la media

$$g(X) = (X - E(x))^r$$

VARIANZA : mide la dispersión promedio de los valores X respecto $E(x)$, desviación respecto de la media al cuadrado

$$\boxed{\sigma_x^2 = E(x - E(x))^2 = E(x^2) - E(x)^2}$$

$$\sigma_{ax+b}^2 = a^2 \sigma_x^2$$

ENTROPIA: mide la incertidumbre en la distribución de la probabilidad en un conjunto. ↑H: probabilidad distribuida más uniforme

$$\boxed{H(X) = -\sum_{x \in X(\Omega)} p(x) \log_2 p(x)}$$

$\log_2 x = \dfrac{\log x}{\log 2}$

$H(x) \leq 0$

MODELOS DE DISTRIBUCIONES DISCRETAS:
- dist. Bernouilli
- dist. binomial
- dist. geométrica
- dist. Poisson

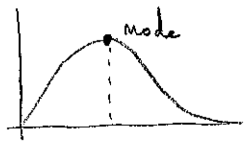
DISTRIBUCIÓN BERNOUILLI: realiza UN experimento aleatorio que puede o no tener éxito

$$\left[\begin{array}{l} P(X=0) = 1-p \quad \text{y} \quad P(X=1) = p \\ \mu = E(x) = p \\ \sigma_x^2 = p(1-p) \end{array}\right]$$

DISTRIBUCIÓN BINOMIAL: mide el número de éxitos en las repeticiones realizadas del experimento

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

$$\boxed{\begin{array}{l} P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \\ \mu = np \qquad \sigma^2 = np(1-p) \end{array}}$$



DIST. GEOMÉTRICA: mide el número de repeticiones hasta que se obtiene el primer éxito del experimento

éxito (uno)
K-1 fracasos } hasta

$$P(x=k) = p\,(1-p)^{k-1}$$

$$\mu = E(x) = \frac{1}{p}$$

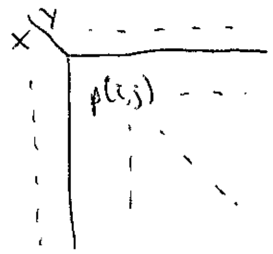
$$\sigma_x^2 = (1-p)/p^2$$

- esta distribución "carece de memoria": en una repetición no influye los resultados anteriores (siempre comienza de cero)

DIST. POISSON: mide un suceso que tiene poca probabilidad de ocurrir en un intervalo de tiempo. Depende de un parámetro $\lambda > 0$

$$\boxed{\begin{gathered} P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!} \\ \mu = E(x) = \lambda \ , \ \sigma_x^2 = \lambda \end{gathered}}$$

VECTOR ALEATORIO: mide simultáneamente varias variables aleatorias en un mismo elemento, sobre un mismo espacio de probabilidad



↓

de forma similar a la función de probabilidad de una única variable

↓

función probabilidad conjunta: $p(i,j) = P(X=i, Y=j)$

$i \in X(\Omega), \ j \in Y(\Omega)$

$p(i,j) \geq 0$

$\sum_i \sum_j p(i,j) = 1$

DISTRIBUCIONES MARGINALES: calculan la probabilidad de una variable aleatoria mientras la otra toma cualquier valor

↓

o distribución unidimensional

$$\boxed{\begin{gathered} P(X=i) = \sum_{j \in Y(\Omega)} P(X=i, Y=j) \\ P(Y=j) = \sum_{i \in X(\Omega)} P(X=i, Y=j) \end{gathered}} \quad \begin{gathered} i \in X(\Omega) \\ j \in Y(\Omega) \end{gathered}$$

- valor esperado de una dist. conjunta: $\boxed{E(f(X,Y)) = \sum_i \sum_j f(i,j)\, P(X=i, Y=j)}$

- valor esperado suma de variables: $E(X_1 + X_2) = E(X_1) + E(X_2)$

- covarianza de dos variables: $\boxed{\rho_{xy} = \frac{\sigma_{x,y}}{\sqrt{\sigma_x^2 \sigma_y^2}}}$ mide la variación simultánea de X e Y

(mirar pág 79 libro)

- dist. condicionada: $\boxed{P(X=x \mid Y=y) = \dfrac{P(X=x, Y=y)}{P(Y=y)}}$ → $P(X=x \mid Y) = P(X=x \mid Y=y_1) + \dots + P(X=x \mid Y=y_n)$

VARIABLES ALETORIAS INDEPENDIENTES: ello significa que el valor de una de ellas no afecta a la distribución marginal de la otra. Propiedad simétrica

$P(X=i, Y=j) = P(X=i)\,P(Y=j)$

- extensible a dos sucesos que se distribuyen según esas variables aleatorias
  $P(X\in A, Y\in B) = P(X\in A)\,P(Y\in B)$
- también a dos funciones de variables independientes: $f(X)$, $g(Y)$
- esperanza de variables independientes: $E(XY) = E(X)E(Y)$
- varianza de la suma de variables independientes: $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$

2. MODELOS PROBABILÍSTICOS CONTINUOS

FUNCIÓN DE DENSIDAD: una función que cumple las condiciones

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

] es función de densidad de probabilidad ≠ función de probabilidad

- es una función de la distribución de la probabilidad en $\mathbb{R}$
- una VARIABLE ALEATORIA CONTINUA es aquella cuya distribución es sobre una función de densidad

$$P(x\in I) = \int_I f(x)\,dx \quad ó \quad P(a<x<b) = \int_a^b f(x)\,dx$$

VALOR ESPERADO VARIABLE CONTINUA: es el valor esperado de una variable continua X sobre una función de densidad (promedio de los valores que toma y la probabilidad en esos valores)

$$\boxed{E(x) = \int_{-\infty}^{\infty} f(x)\,dx}$$

↓
si $E(x) = \infty$, la variable no tiene valor medio

VARIANZA DE UNA VARIABLE CONTINUA: $\boxed{\sigma_x^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - E(x)^2}$

$$\underbrace{\int_{-\infty}^{\infty} x^2 f(x)\,dx}_{E(x^2)} - E(x)^2$$

MODELOS FUNCIONES DENSIDAD:
- función densidad uniforme
- función densidad exponencial
- función densidad normal

FUNCIÓN DENSIDAD UNIFORME: elige un punto al azar en un intervalo $[a,b]$ con $a<b$

$$E(x) = \frac{a+b}{2}$$

$$\sigma_x^2 = \frac{1}{12}(b-a)^2$$

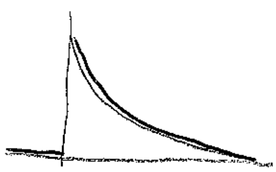
$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a,b] \\ 0, & \text{resto} \end{cases}$$

es una función constante en el intervalo

FUNCIÓN DENSIDAD EXPONENCIAL: depende de un $\lambda > 0$ que hace disminuir la función cuanto más alto sea

$$E(x) = \frac{1}{\lambda} \quad ; \quad \sigma_x^2 = \frac{1}{\lambda^2}$$

$$f(x) = \begin{cases} 0, & \text{si } x < 0 \\ \lambda e^{-\lambda x}, & \text{si } x \geq 0 \end{cases}$$



son similares a la dist. geométrica y "carecen de memoria"

FUNCIÓN DENSIDAD NORMAL: depende de 2 parámetros μ y σ donde

$\mu \to$ media distribución

$\sigma^2 \to$ varianza distribución

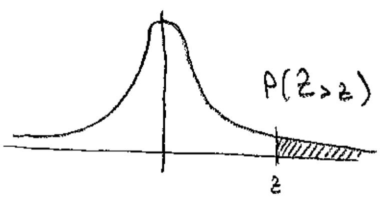
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \Leftarrow \quad f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$$

$\downarrow$

$X - N(0,1)$: simétrica y tabulada

$\downarrow$

$X - N(\mu,\sigma) \to X$ tiene una distribución normal de parámetros μ, σ



- transformación lineal de una variable normal: $E(ax+b) = a E(x) + b = a\mu + b$

$\sigma^2_{ax+b} = \sigma^2_{ax} = a^2\sigma^2$

CALCULOS CON LA DENSIDAD NORMAL: cuando $X \sim N(\mu,\sigma)$, la transformación $\boxed{X \to \frac{x-\mu}{\sigma} = z}$ tiene distribución $N(0,1)$

$P(N(\mu,\sigma) > x) = P(N(0,1) > z)$

$\downarrow$

$z = \frac{x-\mu}{\sigma}$

↳ tipificación de X ejemplo pág 112

- $P(z < -z) = P(z > z)$, $z > 0$
- $P(z \geq z) = P(z > z)$ → valores tabulados $N(0,1)$
- $P(z \leq z) = P(z < z)$
  ↳ $= 1 - P(z > z)$

- $P(0 < x < a) = P(x > 0) - P(x > a)$
- $P(b < x < a) = P(x > b) - P(x > a)$
- $P(-b < x < a) = 1 - P(x > b) - P(x > a)$

↳ cálculo por complementario

guiarse por gráfica también

FUNCIÓN DE DISTRIBUCIÓN: en cada punto, devuelve la probabilidad de que la variable tiene valores menores o iguales al de ese punto

↓
describe cómo se reparte la probabilidad entre los valores de X

$F(x) = P(X \leq x)$, $x \in \mathbb{R}$

3 condiciones: 1. F no decreciente: $x < x'$, $F(x) < F(x')$

2. F continua por la derecha

3. $\lim_{x\to\infty} F(x) = 1$ y $\lim_{x\to-\infty} F(x) = 0$

- puede tener intervalos no disjuntos, todo los puntos en X tienen valor en F(x)
- $[a,b]$: $F(b) - F(a^-) = P(a \leq x \leq b)$
- (a,b): $F(b^-) - F(a) = P(a < x < b)$
- $[a,b)$: $F(b^-) - F(a^-) = P(a \leq x < b)$
- $P(x=a) = P(x \leq a) - P(x < a) = F(a) - F(a^-)$
  ↳ prob. acumulada en un punto

pág 116

→ especialmente en funciones de distribución de variables discretas

- en una función de distribución de una variable discreta hay tantos saltos como valores de X y el tamaño de los saltos es la probabilidad concentrada en esos puntos (fuera de los saltos no hay probabilidad en el intervalo)

- en una función de distribución de una variable continua se tiene que:

$$\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx}$$

↓

si $\int_{-\infty}^{\infty} F'(x)\,dx = 1$ entonces X es variable aleatoria continua

FUNCIÓN DENSIDAD CONJUNTA: se define sobre un vector aleatorio (X, Y) con una distribución de probabilidad continua $f(x,y)$

$$P(X \in I_1, Y \in I_2) = \iint_{I_1 \times I_2} f(x,y)\,dx\,dy$$

- $f(x,y)$ es función densidad conjunta si: $\iint f(x,y)\,dy\,dx = 1$

$f(x,y) \geq 0$ para $-\infty < x < \infty$, $-\infty < y < \infty$

- funciones densidad marginales:

$$\boxed{\begin{aligned} f_x(x) &= \int_{-\infty}^{\infty} f(x,y)\,dy \\ f_y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \end{aligned}}$$

→ x fijo, y invariable

- funciones densidad condicionadas:

$$\boxed{\begin{aligned} \underset{(Y|X=x)}{f_{Y|X}(y)} &= \begin{cases} \dfrac{f(x,y)}{f_x(x)}, & \text{si } f_x(x) > 0 \\ 0, & \text{si } f_x(x) = 0 \end{cases} \\ \underset{(X|Y=y)}{f_{X|Y}(x)} &= \begin{cases} \dfrac{f(x,y)}{f_y(y)}, & \text{si } f_y(y) > 0 \\ 0, & \text{si } f_y(y) = 0 \end{cases} \end{aligned}}$$

↓

modelo dinámico:

$$f(x,y) = f_x(x)\,f(y|x) = f_y(y)\,f(x|y)$$

- independencia de variables: $f(x,y) = f_x(x)\, f_y(y)$ para cada (x,y)
  ↳ solo hay independencia en recintos rectangulares
- valores esperados: $E(x) = \int_{-\infty}^{\infty} x\, f_x(x)\, dx$; $E(y) = \int_{-\infty}^{\infty} y\, f_y(y)\, dy$

3. MUESTREO ALEATORIO

INFERENCIA ESTADÍSTICA: a partir de observaciones repetidas de un fenómeno aleatorio se puede reducir la incertidumbre en algún aspecto de la distribución de probabilidad que lo rige y tomarlo como verdadero para su estudio

↓

hace análisis y elabora previsiones en base a funciones calculadas a partir de los datos muestrales

↳ MUESTRA: datos obtenidos de la observación
↳ DIST. POBLACIÓN: distrib. de la variable observada

MUESTRA ALEATORIA SIMPLE: son las observaciones realizadas de una variable aleatoria $X = (x_1, x_2, \dots, x_n) \rightarrow n$: tamaño de la muestra

↓

ESTADÍSTICO: función de los valores muestrales

↓

DISTRIBUCIÓN MUESTRAL: distribución de la probabilidad de los resultados posibles del muestreo realizado

- si X discreta: $p(x_1, x_2, \dots, x_n) = p(x_1)\, p(x_2) \dots p(x_n)$
- si X continua: $f(x_1, x_2, \dots, x_n) = f(x_1)\, f(x_2) \dots f(x_n)$

- cuanto mayor es n, el estadístico media muestral más se aproxima a la media de la distribución

$$X - \text{dist.}(\mu, \sigma) \Rightarrow \overline{X} - \text{dist}\left(\mu, \frac{\sigma^2}{n}\right)$$

DISTRIBUCIONES PROPIAS DEL MUESTREO: - dist. χ^2 Pearson
(generalmente en poblaciones normales) - dist. t Student

DIST. χ^2 PEARSON: sobre una población teórica $N(0,1)$, una var. aleatoria discreta X de tamaño n la muestra, la suma de cuadrados

$Y = \sum x_i^2$ tiene una distribución χ^2 con n grados de libertad (nº sumandos que aportan variabilidad)

$$E(Y) = n \qquad \sigma_y^2 = 2n$$

- si X tiene una distribución $N(0,\sigma)$: $X - N(0,\sigma)$, el estadístico

$\boxed{\frac{1}{\sigma^2} \sum_{i=1}^{n} x_i^2}$ tiene una dist. χ_n^2

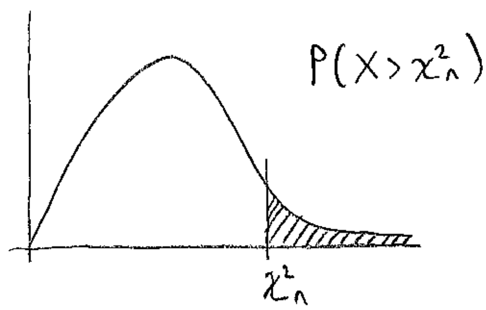
$$\boxed{\begin{aligned} \bar{X} &= \frac{1}{n} \sum x_i && \text{media muestral} \\ s^2 &= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{X}^2 && \text{varianza muestral} \\ S^2 &= \frac{n}{n-1} s^2 && \text{cuasivarianza} \end{aligned}}$$

s^2 y $\bar{X}$ son independientes y tienen una dist. $\chi_{n-1}^2 \longleftarrow$

↓

$n-1$ grados libertad

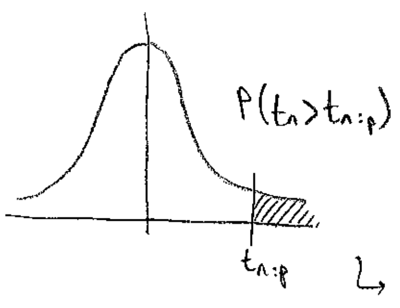
- la dist. χ^2 tiene valores tabulados para distintos grados de libertad ↓tabla y valores de probabilidad → tabla



DIST. t STUDENT: sobre una muestra aleatoria simple $(X_1, X_2, \ldots X_n)$ de una población $N(\mu,\sigma)$, el estadístico t Student:

$$\boxed{t = \sqrt{n-1}\, \frac{\bar{X} - \mu}{s}}$$

tiene dist. t Student con $n-1$ grados de libertad



↳ como es simétrica respecto origen, tiene mismas propiedades de cálculo de $N(0,1)$

4. INFERENCIA ESTADISTICA

ESTIMACIONES POR PUNTO: un estimador puntual es un estadístico función de las observaciones muestrales de una variable aleatoria simple X que permite tomar decisiones sobre el parámetro deseado de la distribución

$$\theta : \hat{\theta} = T(X_1, X_2, \dots, X_n)$$

$$\hat{\theta} \equiv \text{valor estimado de } \theta$$

- lo más deseable a un estimador es que su valor esperado coincida con el parámetro a estimar: $E(\hat{\theta}) = \theta$
- La diferencia entre ellos se denomina SESGO: $b = E(\hat{\theta}) - \theta$

ESTIMADOR CENTRADO: es aquel en que el sesgo es 0
↳ o insesgado

2 estimadores insesgados: media muestral y cuasivarianza muestral

$$\boxed{\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \qquad S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)}$$

1 estimador sesgado: varianza muestral $\hat{s}^2 \to \sigma^2$

$$\boxed{s^2 = \frac{1}{n}\left(\sum x_i^2\right) - \bar{X}^2} \qquad \text{con sesgo } -\sigma^2/n$$

ESTIMADOR SUFICIENTE: un estimador es suficiente para la estimación de un parámetro si la dist. de la muestra condicionada por el estimador es independiente del parámetro

ESTIMADOR MÁXIMA VEROSIMILITUD: dada una función de densidad de una muestra $f(x_1, x_2, \dots, x_n : \theta)$ el estimador $\hat{\theta}$ es de máxima verosimilitud si f alcanza su máximo cuando $\theta = \hat{\theta}$

$$f(x_1, \dots x_n : \hat{\theta}) = \max f(x_1, \dots, x_n : \theta)$$

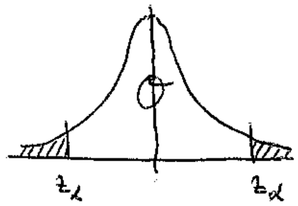
- el principio de máx. verosimilitud establece que la aparición de un suceso atribuido a un modelo entre los modelos probabilísticos del fenómeno hace máxima la probabilidad de que ocurra el suceso
- es aplicable a cualquier aplicación biyectiva de la función de este

INTERVALOS DE CONFIANZA: la estimación por intervalos de confianza consiste en encontrar un intervalo en el que con una probabilidad estimada, se encuentre el estadístico Θ a partir del $\hat{\Theta}$ obtenido en el muestreo

márgenes de variación alrededor de $\hat{\Theta}$ donde Θ se encuentra ←

- Dado un nivel de confianza se encontrará un valor z tal que:

$$P(-z_\alpha < \Theta < z_\alpha) = \alpha$$



- a mayor amplitud intervalo, mayor imprecisión
- α, n (tamaño muestra) fijos: mayor σ, mayor amplitud
- σ, n fijos: $\uparrow \alpha \rightarrow \uparrow$ amplitud
- α, σ fijos: $\uparrow n \rightarrow \uparrow$ amplitud

- Método de la cantidad pivotal: consiste en, mediante transformaciones de las inecuaciones del intervalo, hacer que el intervalo de confianza no se centre en $T(x_1, \dots x_n : \theta)$ sino en una función del propio θ ($g(\theta)$) con nivel de confianza α

$$P(c_1 \leq T(x_1, \dots, x_n ; \theta) \leq c_2) \geq \alpha$$

$$\downarrow$$

$$P(t_1(x_1, \dots x_n ; c_1) \leq g(\theta) \leq t_2(x_1, \dots x_n ; c_2)) \geq \alpha \quad \leftarrow$$

INTERVALOS DE CONFIANZA EN DISTRIBUCIONES NORMALES: para poblaciones $N(\mu,\sigma)$ y $(X_1,\ldots X_n)$

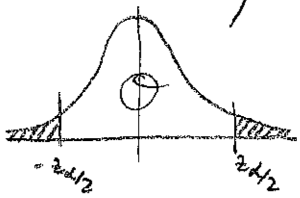
- con σ conocida: para la media μ

$$\boxed{\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)} \quad \text{con} \quad P(z > z_\alpha) = \frac{\alpha}{2}$$

el intervalo se centrará en μ

$z \sim N(0,1)$

$$z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$



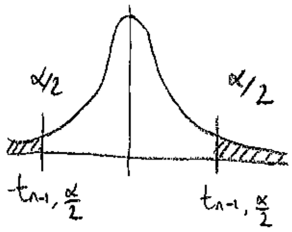
- con σ desconocida: para la media μ

$$\boxed{\left(\bar{X} - t_{n-1;\alpha}\frac{S}{\sqrt{n}},\ \bar{X}+t_{n-1;\alpha}\frac{S}{\sqrt{n}}\right)} \quad \text{con} \quad P(t_{n-1} > t_{n-1,\alpha}) = \frac{\alpha}{2}$$

el intervalo queda centrado en μ

$$t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

$S'^2 \equiv$ cuasivarianza muestra



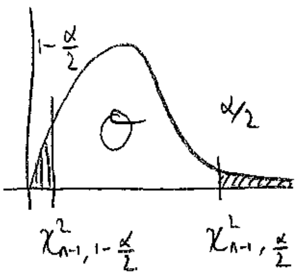
con μ también desconocida

- intervalo para σ:

$$\boxed{\left(\frac{ns^2}{\chi^2_{n-1,\frac{\alpha}{2}}},\ \frac{ns^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}\right)} \quad \text{con} \quad P(\chi^2 > \chi^2_{n-1,1-\frac{\alpha}{2}}) = 1-\frac{\alpha}{2}$$

$$\text{y} \quad P(\chi^2 > \chi^2_{n-1,\frac{\alpha}{2}}) = \frac{\alpha}{2}$$

intervalo centrado en σ



CONTRASTE DE HIPÓTESIS: son técnicas de inferencia que buscan descartar conjeturas acerca de un modelo probabilístico a partir de la información obtenida por la muestra

generalmente:

$H_0 : \theta \in \Theta_0$

$H_1 : \theta \in \Theta_1$

Θ_0, Θ_1 disjuntos

↳ H_0 : hipótesis nula → hipótesis de partida

↳ H_1: hipótesis alternativa → cuando H_0 se rechaza

- para el contraste de hipótesis se realizan pruebas, tests
- la REGIÓN CRÍTICA de un test es el conjunto de muestras que llevan a rechazar H_0 → la REGIÓN de ACEPTACIÓN llevan a aceptar H_0
- como consecuencia de la decisión en el test de hipótesis se pueden dar 2 errores:
  1. Error tipo I → $\boxed{ET1 = P(\text{rechazar } H_0 \mid H_0 \text{ cierta})}$
  2. Error tipo II → $\boxed{ET2 = P(\text{no rechazar } H_0 \mid H_0 \text{ falsa})}$

↓

en el diseño de tests se fija, en función de hipótesis, una cota de probabilidad para cometer ET1 → α

↳ entre los tests que $ET1 < \alpha$, el test que hace ET2 mínima

- la POTENCIA DEL TEST es la probabilidad de rechazar H_0 cuando el valor de θ es θ_1:

$$\boxed{\beta(\theta_1) = 1 - ET2 = P(C \mid \theta = \theta_1)}$$

contrastes paramétricos

CONTRASTES DE HIPÓTESIS UNILATERALES: 4 casos

1. $H_0: \theta = \theta_0$ frente $H_1: \theta > \theta_0$
2. $H_0: \theta \leq \theta_0$ frente $H_1: \theta > \theta_0$

] con $N(\mu, \sigma)$, RC está definida por $\bar{X} > c$

3. $H_0: \theta = \theta_0$ frente $H_1: \theta < \theta_0$
4. $H_0: \theta \geq \theta_0$ frente $H_1: \theta < \theta_0$

] con $N(\mu, \sigma)$, RC está definida por $\bar{X} < c$

CONTRASTE HIPÓTESIS BILATERALES: 1 caso

5. $H_0: \theta = \theta_0$ frente $H_1: \theta \neq \theta_0$ → con $N(\mu, \sigma)$, RC definida sobre $|\bar{X}| \geq c$

CONTRASTE DE BONDAD DEL AJUSTE: verifican si la distribución de la población se reparte ajustada a un patrón permitiendo rechazar la hipótesis que supone que los datos siguen una distribución determinada

contraste no paramétrico

comparando la frecuencia observada con la teórica ←

↳ Ho: probabilidades teóricas

↓

valor discrepancia de Pearson: $$\boxed{D = n\left(-1 + \sum_{i=1}^{n} \frac{\hat{p}_i^2}{p_i}\right)}$$

$\hat{p}_i \equiv$ prob. prácticas (muestra)

$p_i \equiv$ prob. teóricas

$D > 0$ siempre

↓

se distribuye como χ^2 con $n-1$ grados de libertad

Ho rechazada si $D > d^*$ (nivel crítico según nivel significación, α) ←

5. MODELOS DE OPTIMIZACIÓN

MODELO MATEMÁTICO DE OPTIMIZACIÓN: es una representación matemática de un sistema real con 3 elementos

- VARIABLES: representan las alternativas del sistema. Son nº reales
- RESTRICCIONES: igualdades o desigualdades que ligan las variables y representan las relaciones y condiciones del sistema
- FUNCIÓN OBJETIVO: permite comparar alternativas en el conjunto de valores de las variables cumpliendo las condiciones del sistema.
  ↳ debe ser única con el objetivo de maximizar o minimizar el problema
- las variables pueden ser o no controlables según se tomen dentro de las limitaciones del modelo o vengan implícitas de fuera

- Las restricciones relacionan las variables, constantes y parámetros del modelo
  - ↳ pueden venir en la definición del modelo, empíricas, normativas (según el entorno del modelo)...
- Los datos a utilizar han de ser cuantitativos

PROBLEMA DE OPTIMIZACIÓN: consiste en optimizar (encontrar el máximo o el mínimo) de una serie de funciones con n variables y sometido al cumplimiento de unas restricciones

$$\left.\begin{array}{lll} \text{Optimizar} & f(x_1, x_2, \dots, x_n) & \\ \text{sujeto a} & g_i(x_1, x_2, \dots x_n) \geq 0 & i=1,\dots,m \\ & h_i(x_1, x_2, \dots x_n) = 0 & j=1\dots p \end{array}\right]$$

→ si las funciones son lineales: PPL → problema de programación lineal

FORMA GENERAL PPL:

$$\begin{array}{ll} \text{Optimizar} & z = c_1 x_1 + \dots + c_n x_n \\ \text{sujeto a} & a_1 x_1 + \dots + a_n x_n \leq b_1 \\ & \vdots \\ & a_{p1} x_1 + \dots + a_{pn} x_n = b_p \\ & x_1, \dots, x_q \geq 0 \\ & x_{q+1} \dots x_n \text{ cualesquiera} \end{array}$$

restricciones → $\equiv$ restricciones

$x_1, \dots, x_q \geq 0$ → $\equiv$ restricciones no negatividad

$x_i \equiv$ variables decisión

$a_{ij} \equiv$ coeficientes tecnológicos

$b_i \equiv$ vector de disponibilidades

$c_i \equiv$ coeficientes de beneficio / costo

$z \equiv$ función objetivo

↓

para poder tratar el PPL se realizan formulaciones equivalentes del problema

- un problema de maximización se convierte en uno de minimización y viceversa

$$\text{máximo} \sum_i^n c_i x_i = - \text{mínimo} \sum_i^n c_i x_i$$

- las variables x_i no pueden ser negativas. Si x_i puede tomar cualquier valor, se descompone en:

$$x_i = x_i^+ - x_i^- \quad \text{con} \quad x_i^+ = \max\{0, x_i\} \geq 0$$
$$x_i^- = \max\{0, -x_i\} \geq 0$$

- las ecuaciones se pueden pasar a inecuaciones:

$$a_i x_1 + \dots + a_n x_n = b_i \rightarrow a_1 x_1 + \dots + a_n x_n \leq b_i$$
$$a_1 x_1 + \dots + a_n x_n \geq b_i \rightarrow -a_1 x_1 - \dots - a_n x_n \leq -b_i$$

- las inecuaciones se pueden convertir en igualdades incluyendo una variable denominada VARIABLE DE HOLGURA

$$a_1 x_1 + \dots + a_n x_n \leq b_i \rightarrow a_1 x_1 + \dots + a_n x_n + s_i = b_i$$
$$a_1 x_1 + \dots + a_n x_n \geq b_i \rightarrow a_1 x_1 + \dots + a_n x_n - s_i = b_i$$

FORMA CANÓNICA PPL:

maximizar: $z = c_1 x_1 + \dots + c_n x_n$

sujeto a: $a_1 x_1 + \dots + a_n x_n \leq b_1$

$\vdots$

$a_{m1} x_1 + \dots + a_{mn} x_n \leq b_m$

$x_1, \dots, x_n \geq 0$

FORMA STANDARD PPL:

maximizar: $z = c_1 x_1 + \dots + c_n x_n$

sujeto a: $a_1 x_1 + \dots + a_n x_n = b_1$

$\vdots$

$a_{m1} x_1 + \dots + a_{mn} x_n = b_m$

$x_1, \dots, x_n \geq 0$

- Los modelos a su vez pueden expresarse en forma matricial:

maximizar: $z = cx$

sujeto a: $Ax \leq b$

$x \geq 0$

maximizar: $z = cx$

sujeto a: $Ax = b$

$x \geq 0$

$$A = \begin{pmatrix} a_{11} & \cdots & \\ \vdots & \ddots & \\ \vdots & & a_{mn} \end{pmatrix} ; \quad c = (c_1, \dots, c_n) ; \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} ; \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- el PROGRAMA o SOLUCIÓN FACTIBLE del PPL son los valores de x que satisfacen las restricciones del problema → todas las variables son ≥0

↳ su conjunto es la REGIÓN FACTIBLE
↓ es un polígono convexo son los puntos del mismo

- los VERTICES son puntos de la región factible que corresponden con los valores de un programa básico
→ unen ARISTAS

- las soluciones factibles que están sobre una arista forman la FRONTERA y las de dentro de la región factible, INTERIOR

- el PROGRAMA ÓPTIMO es el programa de PPL que hace que la función objetivo alcance su valor máximo

↳ REGIÓN ÓPTIMA

↳ VALOR ÓPTIMO: el de la función objetivo en un programa óptimo

- si un PPL tiene más de un programa óptimo, tiene infinitas soluciones óptimas

- si la región factible no está acotada puede suceder:

↳ tiene un valor óptimo infinito

↳ " finito

- la región factible puede quedar vacía → el PPL es no factible y no tiene solución óptima cualquiera sea la función objetivo