TEMA XI: INFERÉNCIA ESTATÍSTICA

Inferéncia estatística — é a deducción do modelo de probabilidade que segue unha povoación a partir dunha mostra, xeneralizando.

***** Tipos de Inferéncia *segundo o obxectivo* $\begin{cases} \text{Paramétrica (vemos esta)} \\ \text{Non Paramétrica} \end{cases}$

a) Inferéncia **Paramétrica** — Aplíca-se esta inferéncia cando se coñece o modelo da distribución á que pertence a variábel; hai que averiguar os parámetros que caracterizan a devandita distribución.

b) Inferéncia Non Paramétrica — Aplíca-se esta inferéncia cando non se coñece a distribución á que pertence a variábel a partir de dados e de hipóteses mui xerais. A única información da que se dispón é se a variábel é discreta ou contínua, suave ou con picos, etc.

Estes dous procesos de inferéncia normalmente van xuntos.

a) Inferéncia Paramétrica
- *Estimación Puntual* — pregúnta-se cales son os parámetros da povoación.
- *Estimación por Intervalos* — asígna-se un valor aos parámetros nun intervalo, con un grau de certeza desxado.
- *Contraste de Hipóteses* — supón-se algo relativo á povoación e contrásta-se co que indica a mostra.

b) Inferéncia Non Paramétrica
- *Contraste de Hipóteses Non Paramétricos*

***** Tipos de Inferéncia *segundo a información empregada* $\begin{cases} \text{Clásica (vemos esta)} \\ \text{Bayesana} \end{cases}$

i) **Inferéncia Clásica** — supóñen-se os parámetros cantidades fixas descoñecidas e que a única información que hai deles está na mostra.

ii) Inferéncia Bayesana — os parámetros consideran-se v.a. e póde-se introducir información a priori acerca deles.

Mostreo — é o proceso de recollida de dados mediante a realización dun experimento.
Observacións — resultados obtidos nun experimento.
Mostra — conxunto de todos os resultados obtidos nun experimento.

Hai dous tipos de mostreo:

***** Tipos de Mostreo $\begin{cases} \text{con reemprazamento (binomial)} \\ \text{sen reemprazamento (hiperxeométrica)} \end{cases}$

Distribucións coas que se traballa

distribución povoacional	a teórica (p.ex. $x \in N(\mu,\sigma)$, supoñémo-la)
distribución mostral	a prática (p.ex. $x=(x_1,x_2,...,x_{100})$ a marxinal de $x_i \in N(\mu,\sigma)$)
distribución do estimador	para mostras distintas (p.ex. a que seguen as distintas médias $\bar{x}$ tomándo-as como estimadores) A distribución dos $\bar{x}$ depende da povoacional e ainda que $x \notin N(\mu,\sigma)$, o estimador seria aproximadamente normal polo TCL (p.ex. o sumatório de 100 números).
distribcuión empírica	a asociada a cada x_i da mostra

Teorema Fundamental da Estatística

Cando o tamaño mostral n tende a infinito a mostra tende á povoación e polo tanto a distribución empírica tende á distribución povoacional ou teórica.

Glosário

Mostreo Aleatório Simple é o realizado de tal forma que todos os elementos da povoación teñen a mesma probabilidade de resultar-en escolleitos. Realíza-se con reemprazamento, a povoación é a mesma en cada unha das extraccións.

Mostra Aleatória Simple Sexa x unha v.a., unha mostra aleatória simple (m.a.s.) de x é un vector aleatório $(x_1,x_2,...,x_n)$ tal que x_i ten a mesma distribución que x e son independentes.

Estatístico Sexa $(x_1,x_2,...,x_n)$ unha m.a.s. de x —as x_i son variábeis aleatórias—; un *estatístico* T é unha variábel aleatória. chamada *estatístico* se $T=H(x_1,x_2,...,x_n)$, onde H é unha función.

Estimador é un estatístico (unha variábel aleatória) que se utiliza para estimar un parámetro. Unha estimación é un número que representa a avaliación do estimador nunha realización concreta da mostra.

Características Mostrais

μ	**Média mostral**	$\bar{x}=\sum_{i=1}^{n} \frac{x_i}{n}$
σ^2	**Varianza mostral**	$S_n^2=\sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{n}$
$\alpha_k=E(x^k)$	**Momentos mostrais**	$\alpha_k=\sum_{i=1}^{n} \frac{x_i^k}{n}$

Outros tipos de Mostreo

[1] **Mostreo Sistemático** tóma-se o primeiro elemento de forma aleatória e logo tóman-se elementos a unha distáncia fixa que se despraza para ambos lados a partir do primeiro: ...←16←20→24→...

[2] **Mostreo Estratificado** considéra-se que a povoación non é homoxénea senón que está dividida en estratos. En cada estrato realíza-se o mostreo por mostreo aleatório simple e o número de elementos seleccionados en cada estrato sóe-se tomar na mesma proporción que existe entre o número de individuos dese estrato respeito da povoación total.

[3] **Mostreo Polietápico** Aquel mostreo que se realiza nunha rexión. Hai diferentes etapas. P.ex. unha mostra de persoas, povoación complexa, fai-se por bairros, ruas, edificios...

[4] **Mostreo por Conglomerado** Se hai cinco provincias co mesmo comportamento, só se toma unha. Contrário a estratificado, cando a mostra se comporta igual que a povoación.

[5] **Simulación** Simulo dados ⇒ coñezo o seu modelo ⇒ uso un método e contrasto resultados co modelo. O método será válido se hai coincidéncia.

Xeralización de variáveis aleatórias (mostras aleatórias simples dunha povoación cuia distribución se descoñece)
Método de Montecarlo (para xerar unha mostra de tamaño n dunha povoación con F(x) coñecida)

1º) xerar $U_1 \in [0,1]$ número aleatório con tantos decimais como se desexe.
2º) $U_1=F(x) \Rightarrow$ tóma-se como valor observado $x_1=F^{-1}(U_1)$
3º) repetir o proceso n veces

A partir da xeración de números aleatórios (que seguen a distribución U(0,1)):

uniforme	1) se U v.a. $\in U(0,1) \Rightarrow x=U(b-a)+a \in U(a,b)$
binomial	2) xéran-se n v.a. de Bernouilli $y_1,y_2,...,y_n$ (y_i=1 con probabilidade p, 0 con 1-p) entón $x= y_1+y_2+...+y_n \in B(n,p)$
expoñencial	3) se U v.a. $\in U(0,1) \Rightarrow x=1/\lambda \cdot \ln(1/(1-U)) \in Exp(\lambda)$
normal	4) U_1, U_2 números aleatórios $\Rightarrow \begin{cases} x_1=\sqrt{(-2\cdot\ln U_1)}\cdot \mathrm{sen}(2\pi U_2) \\ x_2=\sqrt{(-2\cdot\ln U_1)}\cdot\cos(2\pi U_2) \text{ son v.a.indep.} \in N(0,1) \end{cases}$
(a partir de 4)	* 5) X^2,F,t trivial
	* 4) tamén polo TCL: n v.a. $U(0,1) \Rightarrow x=U_1+ U_2+...+ U_n-6 \approx N(0,1)$