TEMA XIII: MÉTODOS DE OBTENCIÓN DE ESTIMADORES

• **Método dos Momentos** (Pearson) consiste en igualar os momentos teóricos aos momentos mostrais en número igual aos parámetros a estimar. Así obten-se un sistema de ecuacións cuias incógnitas son os parámetros.

Os estimadores por momentos son consistentes (o mínimo que se pode esixir) pero non son centrados nen eficientes (non teñen mínima varianza). Avantaxe: sinxeleza, val para a primeira aproximación. Inconvinte: ignora a distribución da povoación á que pertence a mostra; non usa toda a información da que dispón.

• **Método de Máxima Verosimilitude** (Fisher) 1º) $P(\vec{x}/\vec{\theta})$ 2º) $\vec{\theta}/P(\vec{x}/\vec{\theta})$ max.

Sexa x unha v.a. con función $f(x,\vec{\theta})$ —f probabilidade se x discreta, f densidade se contínua—, e $\vec{\theta}$ vector de parámetros descoñecido do que depende a función de probabilidade/densidade. Se se realiza un mostreo aleatório simple, a mostra $\vec{x}=(x_1...x_n)$ é unha v.a. n-dimensional con densidade conxunta $f(\vec{x},\vec{\theta})=\Pi_{i=1}^{n} f(x_i,\vec{\theta})$, onde x_i coñecido, $\vec{\theta}$ descoñecido. As observacións mostrais x_i son v.a.i.i.d. con función (prob./dens.) $f(x,\vec{\theta})$. Definimos a función de verosimilitude como a probabilidade de obter $\vec{x}$ cando $\vec{\theta}$ ten un valor determinado:

$$l(\vec{\theta})=P(\vec{x}/\vec{\theta})\underset{\text{indep.}}{\overset{}{=}}P(x_1/\vec{\theta})\cdot P(x_2/\vec{\theta})\cdot...\cdot P(x_n/\vec{\theta})=\Pi_{i=1}^{n} f(x_i,\vec{\theta})$$

O método de máxima verosimilitude asigna a $\vec{\theta}$ o valor tal que $l(\vec{\theta})$ é máximo —se $l(\theta_1)>l(\theta_2)$, θ_1 é mais probábel que θ_2 xa que $P(\vec{x}/\theta_1)>P(\vec{x}/\theta_2)$—.

Para evitar as unidades: $l(\theta_1)/l(\theta_2)>1\Rightarrow \theta_1$ mais probábel que θ_2
$l(\theta_1)/l(\theta_2)>1\Leftrightarrow \ln l(\theta_1)>\ln l(\theta_2)$

Definimos a función suporte $L(\vec{\theta})=\ln l(\vec{\theta})$

ousexa que se $L(\theta_1)>L(\theta_2)$ entón θ_1é mais probábel que θ_2

Selecciónа-se como estimador do parámetro o valor que maximice a probabilidade do que ocurriu (a mostra). Derívа-se a función suporte para cada $\theta_i \in \vec{\theta}$ e iguála-se a cero, obtendo os estimadores máximo verosímiles.

Propriedades dos estimadores máximo-verosímiles

[1] Asintoticamente Centrados $\lim_{n\to\infty} E(\bar{\theta}_{mv})=\theta$

[2] Distribución Asintoticamente Normal $\bar{\theta}_{mv}\approx N$ cando $n\to\infty$

[3] Asintoticamente Eficientes/de Varianza Minima $\lim_{n\to\infty} ECM(\bar{\theta}_{mv})=0$

[4] $\bar{\theta}_{mv}$=estimador MV do parámetro $\theta \Rightarrow g(\bar{\theta}_{mv})$=Estimador MV parámetro $g(\theta)$ (prop. mui prática)

[5] Existe estatístico suficiente* para $\theta \Rightarrow \bar{\theta}_{mv}$ é suficiente

*estatístico suficiente é o que usa toda a información que a mostra dá sobre o parámetro a estimar. Por exemplo, estimo μ por $\bar{x}$, $\bar{x}$ é un estatístico suficiente, non é necesário coñecer os x_i da mostra.

Estimación dos Parámetros en Povoacións Normais

Os estimadores de média e varianza povoacional son valores mostrais.

Método dos Momentos: $\hat{\mu}_{mom}=\bar{x}=1/n\Sigma_{i=1}^{n}x_i$ $\hat{\sigma}^2_{mom}=s^2=1/n\Sigma_{i=1}^{n}(x_i-\bar{x})^2$

Método de Máxima Verosimilitude

$f(x,\vec{\theta})=1/\sqrt{(2\pi\sigma^2)}\cdot\exp\{-1/2\cdot(x-\mu)^2/\sigma^2\} \Rightarrow l(\vec{\theta})=\Pi_{i=1}^{n}1/\sqrt{(2\pi\sigma^2)}\cdot\exp\{-1/2\cdot[(x_i-\mu)/\sigma)]^2\}=$
$=1/\sqrt{(2\pi\sigma^2)}\cdot\exp\{\Sigma_{i=1}^{n}-1/2\cdot[(x_i-\mu)/\sigma)]^2\}=1/\sqrt{(2\pi)}\cdot n\cdot 1/\sqrt{(v^n)}\cdot\exp\{\Sigma_{i=1}^{n}(x_i-\mu)^2/(-2v)\}$

$L(\vec{\theta})=-n\cdot\ln\sqrt{(2n)}-n\cdot\ln\sqrt{n}+\Sigma_{i=1}^{n}(x_i-\mu)^2/v=-n/2\cdot\ln v-1/(2v)\cdot\Sigma_{i=1}^{n}(x_i-\mu)^2$

$dL(\vec{\theta})/d\mu=-1/(2v)\cdot\Sigma_{i=1}^{n}(2x_i-2\mu)=\Sigma_{i=1}^{n}(\mu-x_i)=n\mu-\Sigma_{i=1}^{n}x_i=0 \Rightarrow$
$\Rightarrow \hat{\mu}_{mv}=1/n\cdot\Sigma x_i=\bar{x}$

$dL(\vec{\theta})/dv=-n/2\cdot 1/v+v^{-2}\cdot 1/2\cdot\Sigma_{i=1}^{n}(x_i-\mu)^2=1/(2v)\cdot(-n+1/v\cdot\Sigma_{i=1}^{n}(x_i-\bar{x})^2)=0 \Rightarrow$
$\Rightarrow n=1/v\cdot\Sigma_{i=1}^{n}(x_i-\bar{x})^2 \Leftrightarrow \hat{v}_{mv}=1/n\cdot\Sigma_{i=1}^{n}(x_i-\bar{x})^2=s^2$

Estimación Paramétrica de Povoacións Normais

• propriedades de $\bar{x}$ (média mostral)

como $\bar{x}$ é unha combinación lineal de v.a. normais, segue unha distribución normal:

$\bar{x}\in N(\mu,\sigma/\sqrt{n})$ xa que $E(\bar{x})=\mu$; $Var(\bar{x})=\sigma^2/n$ $\Rightarrow (\bar{x}-\mu)/(\sigma/\sqrt{n})\in N(0,1)$

↑

exactamente porque a distribución x é N (diciamos que para calquera distribución de x cando $n\to\infty$, a distribución de $\bar{x}$ era aproximadamente N polo TCL)

• proriedades de s^2 (varianza mostral) úsa-se cando $n\to\infty$

como $x_1,...,x_n$ son v.a.i. normais que verifican $\Sigma_{i=1}^{n}(x_i-\bar{x})=0$, e a suma de normais (0,1) ao cuadrado define a Chi Cuadrado, $\Sigma_{i=1}^{n}((x_i-\mu)/\sigma)^2\in X_{n-1}^2$ onde $x_1-\bar{x},...,x_n-\bar{x}$ son n-1 v.a.i., entón:

$$ns^2/\sigma^2\in X_{n-1}^2$$

$E(ns^2/\sigma^2)=E(X_{n-1}^2)=n-1 \Leftrightarrow n/\sigma^2\cdot E(s^2)=n-1 \Leftrightarrow \underline{E(s^2)=(n-1)/n\cdot\sigma^2}$

$Var(ns^2/\sigma^2)=Var(X_{n-1}^2)=2(n-1) \Leftrightarrow n^2/\sigma^4\cdot Var(s^2)=2\cdot(n-1) \Leftrightarrow \underline{Var(s^2)=\sigma^4/n^2\cdot 2\cdot(n-1)}$

Acerca da distribución de s^2: como s^2 é unha suma de v.a., s^2 é aproximadamente normal polo TCL.
s^2 é sesgado pero asintoticamente insesgado, por iso definimos unha corrección da varianza mostral que é estimador insesgado de σ^2: a cuasivarianza mostral $\hat{s}^2$

$\underline{\hat{s}^2=}1/(n-1)\cdot\Sigma_{i=1}^{n}(x_i-\bar{x})^2=\underline{n/(n-1)\cdot s^2}$

$E(\hat{s}^2)=E(n/(n-1)\cdot s^2)=n/(n-1)\cdot E(s^2)=n/(n-1)\cdot(n-1)/n\cdot\sigma^2=\sigma^2$ $\Rightarrow$ é un estimador centrado de σ^2

• <u>propriedades de $\hat{s}^2$</u> (a cuasivarianza mostral)

$E(\hat{s}^2)=\sigma^2$
$Var(\hat{s}^2)=2/(n-1)\cdot\sigma^4$ porque $ns^2/\sigma^2 \in X_{n-1}^2 \Rightarrow [n\cdot(n-1)/(n-1)\cdot s^2]/\sigma^2 \in X_{n-1}^2 \Rightarrow (n-1)\hat{s}^2/\sigma^2 \in X_{n-1}^2$
sabemos que $Var((n-1)\hat{s}^2/\sigma^2)=Var(X_{n-1}^2)=2(n-1)=(n-1)^2/\sigma^4\cdot Var(\hat{s}^2)$

<u>Comparación s^2 e $\hat{s}^2$</u>

ECM(s^2)$=E(s^2-\sigma^2)=(\text{sesgo } s^2)^2+Var(s^2)=(E(s^2)-\sigma^2)^2+Var(s^2)=((n-1)/n\cdot\sigma^2-\sigma^2)^2+\sigma^4/n^2\cdot 2(n-1)=$
$=\sigma^4\cdot[((n-1)/n-n/n)^2+2/n^2\cdot(n-1)]=\sigma^4\cdot(1/n^2+(2n-2)/n^2)=$ $\boldsymbol{\sigma^4\cdot(2n-1)/n^2}$

ECM($\hat{s}^2$)$=E(\hat{s}^2-\sigma^2)^2=(\text{sesgo } \hat{s}^2)^2+Var(\hat{s}^2)=(E(\hat{s}^2)-\sigma^2)^2+Var(\hat{s}^2)=(\sigma^2-\sigma^2)^2+2/(n-1)\cdot\sigma^4=$ $\boldsymbol{2/(n-1)\cdot\sigma^4}$

$ECM(s^2)>ECM(\hat{s}^2)$// $\sigma^4\cdot(2n-1)/n^2>2/(n-1)\cdot\sigma^4$//$2n^2-n-2n+1>2n^2$//$+3n>+1$//$n>1/3$

Conclusións: s^2 ten menos erro cuadrático médio que $\hat{s}^2$
$\hat{s}^2$ insesgado
s^2 sesgado
$n\to\infty \Rightarrow s^2\approx\hat{s}^2$