TEORIA DELLA STIMA:
MASSIMA VEROSIMIGLIANZA

ε	$\Rightarrow$	esperimento
X	$\Rightarrow$	v.c. che descrive l'esperimento
$f_X(x,\theta)$	$\Rightarrow$	famiglia di distribuzioni a cui appartiene X
θ	$\Rightarrow$	parametro (mono o pluri dimensionale) che caratterizza la densità di probabilità

Conoscendo $f_X(x,\theta)$ cui X appartiene per $\theta=\bar{\theta}$ ed avendo un campione (bernoulliano) $\{x_1,\dots,x_N\}$ estratto da X vogliamo ottenere un stima $\hat{\theta}$ di $\bar{\theta}$ tale che $|\hat{\theta}-\bar{\theta}|$ sia piccolo, con probabilità alta.

N.B. $\hat{\theta}$ Dipende dai valori del campione, quindi $\hat{\theta}=\hat{\theta}(x_1,\dots,x_N)$

Campione bernoulliano: ottenuto con prove indipendenti tutte uguali tra loro

VARIABILE CAMPIONARIA

Variabile campionaria: nel caso bernoulliano, descrive N ripetizioni uguali e indipendenti di ε, descrive quindi l'esperimento complesso ε^N

X $\Rightarrow$ v.c. che descrive la singola prova dell'esperimento

$\underline{X}^{(N)}=\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$ $\Rightarrow$ v.c. , v.camp. N-dimensionale, con $X_1, X_2, \dots, X_N$ normali

$\underline{x}^o=\begin{bmatrix} x_1^o \\ x_2^o \\ \vdots \\ x_N^o \end{bmatrix}$ $\Rightarrow$ campione bernoulliano, estratto da $\underline{X}^{(N)}$

essendo il campione bernoulliano:

$$E_{\underline{X}}\{\underline{X}^{(N)}\}=\begin{bmatrix} \mu_{X_1} \\ \mu_{X_2} \\ \vdots \\ \mu_{X_N} \end{bmatrix}=\begin{bmatrix} \mu_X \\ \mu_X \\ \vdots \\ \mu_X \end{bmatrix}=\mu_X\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{e} \quad C_{\underline{XX}}=\begin{bmatrix} \sigma_X^2 & 0 & 0 \\ 0 & \sigma_X^2 & 0 \\ 0 & 0 & \sigma_X^2 \end{bmatrix}=\sigma_X^2\cdot I_N$$

$$f_{\underline{X}}(\underline{x};\theta)=f_{\underline{X}}(x_1,\dots,x_N;\theta)=\prod_{i=1}^N f_X(x_i;\theta)$$

$L(\underline{x};\theta)=f_{\underline{X}^{(N)}}(\underline{x};\theta)=\prod_{i=1}^N f_X(x_i;\theta)$ densità di $\underline{X}^{(N)}$, likelihood (verosimiglianza)

STATISTICHE
Si definisce statistica $S(\underline{X}^{(N)})$ una qualsiasi funzione della v.camp. $\underline{X}^{(N)}$
La v.c. *media campionaria* costituisce un esempio di statistica: $M=\frac{1}{N}\sum_{i=1}^{N} X_i$

Dato un campione si può però ottenere una stima della *media campionaria:*
$m=\frac{1}{N}\sum_{i=1}^{N} x_i$ v.s. estrazione dalla v.c. M

STIMATORI
Data una funzione di likelihood $L(\underline{x};\theta)$, trovare una statistica $T(\underline{X}^{(N)})$ distribuita intorno a θ, tale che un'estrazione t da T ($t=T(\underline{x})$), ottenuta da un campione $\underline{x}$ sia vicina a θ con probabilità alta.

$T(\underline{X}^{(N)})$	*stimatore* di θ (statistica)	$\hat{\Theta}$	spesso si indica con $\hat{\theta}$
$t=T(\underline{x})$	*stima* di θ	$\hat{\theta}$	

CORRETTEZZA DEGLI STIMATORI
Se la media dello stimatore coincide con il parametro che si vuole stimare, lo stimatore è *corretto*
Deve essere $E\{T(\underline{X})\}=\theta$

Sia $T(\underline{X}^{(N)})=T(\underline{X})$

$T(\underline{X})$	$\Rightarrow$ distribuito intorno a θ
$b(\theta)=E\{T(\underline{X})\|\theta\}-\theta$	$\Rightarrow$ *bias* (distorsione)
$E\{T(\underline{X})\|\theta\}$	$\Rightarrow$ media di $T(\underline{X})$
θ	$\Rightarrow$ valore corretto del parametro

quindi $E\{T(\underline{X})|\theta\}=\int_{\mathbb{R}^N} T(\underline{x})L(\underline{x};\theta)d_N x$ da $E_X\{g(X)\}=\int_{-\infty}^{+\infty} g(x)f_X(x)dx$

Se $E\{T(\underline{X})\}=\theta$ lo stimatore ha bias nullo e si dice *unbiased* (*corretto*)

CONVERGENZA IN PROBABILITÀ
Una successione converge in probabilità se per $N\to\infty$, $m\to p$ e $s\to 0$ (condizione necessaria)

CONSISTENZA DEGLI STIMATORI
Per essere buono, vogliamo che lo stimatore, quando N tende a infinito, faccia *convergere in probabilità* $T(\underline{X}^{(N)})$ a θ
$\lim_{N\to\infty} T(\underline{X}^{(N)})=\theta$ cioè $\lim_{N\to\infty} P\left[(T(\underline{X}^{(N)})-\theta)<\epsilon\right]=1 \quad \forall\epsilon$

O equivalentemente, applicando le condizioni per la convergenza in media quadratica:

$$\begin{cases} \lim_{N\to\infty} E\{T(\underline{X}^{(N)})\}=\theta & \text{media dello stimatore} \\ \lim_{N\to\infty} \sigma^2\{T(\underline{X}^{(N)})\}=0 & \text{varianza dello stimatore} \end{cases}$$

N.B. Se la media tende a θ , il bias $b(\theta)$ tende a zero per $N\to\infty$

LIKELIHOOD DI VC NOTE

POISSONIANA

$$P(K=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} \qquad (\theta=\lambda)$$

$$L(\underline{K}^{(N)},\lambda)=\prod_{i=1}^{N}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}=\mathrm{e}^{-\lambda N}\frac{\lambda^{\sum_i k_i}}{\prod_{i=1}^{N}k_i!}$$

BINOMIALE

$$P(K=k)=\binom{n}{k}p^k(1-p)^{n-k} \qquad (\theta=p)$$

$$L(\underline{K}^{(N)},p)=\prod_{i=1}^{N}\binom{n}{k_i}p^{k_i}(1-p)^{n-k_i}=\left[\prod_{i=1}^{N}\binom{n}{k_i}\right]p^{\sum_i k_i}(1-p)^{nN-\sum_i k_i}$$

NORMALE

$$f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad (\theta=(\mu,\sigma^2))$$

$$L(\underline{X}^{(N)},(\mu,\sigma^2))=\prod_{i=1}^{N}f_X(x_i)=\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_i(x_i-\mu)^2}$$

ESPONENZIALE

$$f_X(x)=\frac{1}{\mu}\mathrm{e}^{-\frac{x}{\mu}} \qquad (\theta=\mu)$$

$$L(\underline{X}^{(N)},\mu)=\prod_{i=1}^{N}\frac{1}{\mu}\mathrm{e}^{-\frac{x_i}{\mu}}=\frac{1}{\mu^N}\mathrm{e}^{-\frac{\sum_i x_i}{\mu}}$$

STIMATORI DI MINIMA VARIANZA: EFFICIENZA

Uno stimatore si dice efficiente se la sua varianza è più piccola di quella di altri stimatori della stessa variabile

CONVERGENZA IN LEGGE: TEOREMA CENTRALE DELLA STATISTICA

La somma di N v.c. tutte tra loro indipendenti, e con la stessa distribuzione tende asintoticamente in legge ad un normale, qualunque sia la distribuzione iniziale delle N v.c.

Cioè:

Sia $\{X_i\}$ una successione di variabili indipendenti, tutte con la stessa distribuzione e con

$$E\{X_i\}=\mu \quad \text{e} \quad \sigma^2(X_i)=\sigma^2$$

allora la successione $S_n=\sum_{i=1}^{n}X_i$

ha l'andamento asintotico in legge $S_n\sim N[n\mu,n\sigma^2]$

qualunque sia la distribuzione iniziale delle $\{X_i\}$, ovvero $\frac{S_n-n\mu}{\sqrt{n}\sigma}\xrightarrow{L}N[0,1]$

STIME DI MASSIMA VEROSIMIGLIANZA (MAXIMUM LIKELIHOOD)

Si sceglie come stimatore di θ, la statistica $T(\underline{X})$ tale che:

$$L(\underline{x};T(\underline{x}))=\max_{\theta} L(\underline{x};\theta)$$

Dato un campione $\underline{x}$ si sceglie quindi il θ per cui è massima la probabilità di estrarre proprio quell' $\underline{x}$.

ESEMPIO: CAMPIONI NORMALI

$$X\sim N[\mu_X;\sigma_X^2] \quad \Rightarrow \quad f_X(x)=\frac{1}{\sigma_X\sqrt{2\pi}}\mathrm{e}^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}}$$

$$\underline{x}^o=\begin{bmatrix} x_1^o \\ \vdots \\ x_N^o \end{bmatrix} \quad ; \quad \underline{X}^{(N)}=\begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} \quad ; \quad E_{\underline{X}}\{\underline{X}^{(N)}\}=\begin{bmatrix} \mu_X \\ \vdots \\ \mu_X \end{bmatrix}=\mu_X\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad ; \quad C_{\underline{XX}}=\sigma_X^2\cdot I_N \quad ; \quad \theta=\begin{bmatrix} \mu_X \\ \sigma_X^2 \end{bmatrix}$$

$$L(\underline{X}^{(N)},\theta)=L(\underline{X}^{(N)},(\mu,\sigma^2))=\prod_{i=1}^{N} f_X(x_i^{(o)})=\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{1}{2\sigma^2}(x_i^{(o)}-\mu)^2}=\frac{1}{(2\pi)^{\frac{N}{2}}(\sigma^2)^{\frac{N}{2}}}\mathrm{e}^{-\frac{1}{2\sigma^2}\sum_i(x_i^{(o)}-\mu)^2}$$

$$\ln\left(L(\underline{X}^{(N)},\theta)\right)=-\frac{N}{2}\ln(2\pi)-\frac{N}{2}\ln(\sigma^2)-\frac{1}{2}\sum_{i=1}^{N}\frac{(x_i^{(o)}-\mu)^2}{\sigma^2}$$

Per trovare il massimo della funzione deve essere:

$$\begin{cases} \dfrac{\partial\ln(L)}{\partial\mu}=0 \\ \dfrac{\partial\ln(L)}{\partial\sigma^2}=0 \end{cases}$$

$$\frac{\partial\ln(L)}{\partial\mu}=0 \quad \Rightarrow \quad -\frac{1}{\cancel{2}}\sum_{i=1}^{N}\frac{-\cancel{2}(x_i^{(o)}-\mu)}{\sigma^2}=0 \quad \Rightarrow \quad \sum_{i=1}^{N}x_i^{(o)}-\sum_{i=1}^{N}\mu=0 \quad \Rightarrow \quad \sum_{i=1}^{N}x_i^{(o)}-N\mu=0$$

$$\frac{\partial\ln(L)}{\partial\sigma^2}=0 \quad \Rightarrow \quad -\frac{N}{\cancel{2}}\frac{1}{\cancel{\sigma^2}}+\frac{1}{\cancel{2}}\frac{1}{(\sigma^2)^{\cancel{2}}}\sum_{i=1}^{N}(x_i^{(o)}-\mu)^2=0 \quad \Rightarrow \quad -N+\frac{1}{(\sigma^2)}\sum_{i=1}^{N}(x_i^{(o)}-\mu)^2=0$$

$$\begin{cases} \sigma^2=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(x_i^{(o)}-\mu)^2 \\ \displaystyle\sum_{i=1}^{N}x_i^{(o)}-N\mu=0 \end{cases}$$

STIME	STIMATORI	
$\hat{\mu}=\frac{1}{N}\sum_{i=1}^{N}x_i^{(o)}$	$M=\frac{1}{N}\sum_{i=1}^{N}X_i^{(o)}$	media campionaria
$\hat{\sigma^2}=\frac{1}{N}\sum_{i=1}^{N}(x_i^{(o)}-\hat{\mu})^2$	$S^2=\frac{1}{N}\sum_{i=1}^{N}(X_i^{(o)}-M)^2=T(\underline{X}^{(N)})$	varianza campionaria