RICHIAMIAMO ALCUNE COSE SUI VETTORI ALEATORI.

SE $X=(X_1,\dots,X_n)$ HA COMPONENTI DISCRETE CON DENSITÀ $q_1,\dots,q_n$ E SONO INDIPENDENTI, ALLORA X HA DENSITÀ $q(x)=\prod_{k=1}^{n} q_k(x_k)$. IN PARTICOLARE

$$P\{X\in\prod_{k=1}^{n}(a_k,b_k)\}=\prod_{k=1}^{n}\sum_{x_k\in(a_k,b_k)} q_k(x_k)$$

SE LE X_k SONO INVECE ASSOLUTAMENTE CONTINUE CON DENSITÀ $f_1,\dots,f_n$ E SONO INDIPENDENTI, CONTINUA A VALERE QUALCOSA DI SIMILE

$$P\{X\in\prod_{k=1}^{n}(a_k,b_k)\}=\prod_{k=1}^{n}\int_{a_k}^{b_k} f_k(x_k)\,dx_k \qquad (*)$$

LA FUNZIONE $f:\mathbb{R}^n\to[0,+\infty)$ DEFINITA DA $f(x)=\prod_{k=1}^{n} f_k(x_k)$ È CHIAMATA DENSITÀ DI X. IL CONCETTO SI PUÒ ESTENDERE ANCHE AL CASO IN CUI LE COMPONENTI NON SONO INDIPENDENTI, MA IN QUESTO CASO f NON HA LA FORMA DI UN PRODOTTO E IL LATO DESTRO $(*)$ DIVIENE UN INTEGRALE MULTIPLO. NON CI SERVIRÀ NEL SEGUITO E QUINDI SOPRASSEDIAMO.

FACCIAMO UN VELOCE RIEPILOGO

UN MODELLO STATISTICO È UNA FAMIGLIA DI SPAZI DI PROBABILITÀ $\{(\Omega, \mathcal{A}, P_\theta), \theta \in \Theta\}$, DOVE Θ È UN OPPORTUNO INSIEME. PIÙ OPERATIVAMENTE, SI CONSIDERA UNA FAMIGLIA DI LEGGI $\{\mathcal{L}_\theta, \theta \in \Theta\}$ DI V.A. NOTE A MENO DEL PARAMETRO θ.

PER ESEMPIO, VOGLIAMO OPERARE CON UNA V.A. $X: \Omega \to \mathbb{R}$ DI TIPO GAUSSIANO. IL FATTO CHE LA PROBABILITÀ SOTTOSTANTE VARIA CON θ FA SÌ CHE LA LEGGE DI X VARIA

$$X \sim \mathcal{N}(\mu(\theta), \sigma^2(\theta))$$

UN CAMPIONE PER UN MODELLO STATISTICO È UNA SUCCESSIONE DI V.A. $\{X_n\}_{n \in \mathbb{N}}$ I.I.D, CIASCUNA CON LEGGE $\mathcal{L}_\theta$. NEL CASO DISCRETO OGNUNA AVRÀ UNA DENSITÀ $q(x, \theta)$, MENTRE NEL CASO ASSOLUT. CONTINUO UNA DENSITÀ $f(x, \theta)$. L'N-PLA $(X_1, \dots, X_n)$ È DETTA CAMPIONE DI AMPIEZZA n. CHIAMEREMO STATISTICA RELATIVA AL CAMPIONE UNA V.A. DEL TIPO $H = h(X_1, \dots, X_n)$ CON $h: \mathbb{R}^n \to \mathbb{R}$.

Si dice **stimatore** del parametro θ una successione di statistiche $H_n = h_n(X_1, \dots, X_n)$, $h_n : \mathbb{R}^n \to \mathbb{R}$ usate per stimare θ o, più in generale, una sua funzione $\psi(\theta)$. Se $\theta \in \mathbb{R}^m$ in genere si prendono $\psi_1(\theta) = \theta_1, \dots, \psi_m(\theta) = \theta_m$, cioè le proiezioni.

La singola H_n è detta stimatore basato su un campione di taglia n.

Una volta estratto un particolare campione $(x_1, \dots, x_n)$ relativo a $(X_1, \dots, X_n)$, il valore numerico $h(x_1, \dots, x_n)$ è detto **stima** per $\psi(\theta)$. Il fatto di osservare il campione $(x_1, \dots, x_n)$ significa che si è verificato l'evento $\{X_k = x_k, \, k = 1, \dots, n\}$.

Proprietà degli stimatori

Correttezza. Uno stimatore H_n di $\psi(\theta)$ è detto corretto se $E_\theta[H_n] = \psi(\theta)$ $\forall \theta \in \Theta$ e $\forall n \in \mathbb{N}$

Consistenza. Uno stimatore H_n di $\psi(\theta)$ è detto consistente se

$$\lim_{n \to \infty} P_\theta\left(|H_n - \psi(\theta)| > \varepsilon \right) = 0 \quad \forall \varepsilon > 0 \text{ e } \forall n \in \mathbb{N}$$

ABBIAMO VISTO COME LE STATISTICHE

MEDIA CAMPIONARIA $\bar{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$

VARIANZA CAMPIONARIA $S^2_n = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \bar{X}_n)^2$

SIANO STIMATORI CORRETTI E CONSISTENTI, RISPETTIVAMENTE PER LA MEDIA E LA VARIANZA.

INTRODUCIAMO ORA UN ALTRO METODO DI STIMA PUNTUALE: LA <u>MASSIMA VEROSIMIGLIANZA</u>.

SUPPONIAMO CHE LE V.A. $X_1, \dots, X_n$ SIANO DISCRETE O ASS. CONTINUE CON DENSITÀ CONGIUNTA $x \in \mathbb{R}^n \to L(x,\theta)$ (CIOÈ $L(x,\theta) = \prod_{k=1}^{n} q(x_k,\theta)$ NEL CASO DISCRETO, $L(x,\theta) = \prod_{k=1}^{n} f(x_k,\theta)$ NEL CASO ASS. CONTINUO). SUPPONIAMO INOLTRE CHE IL CAMPIONE OSSERVATO SIA $\hat{x} = (x_1, \dots, x_n)$. UN MODO RAGIONEVOLE PER SCEGLIERE IL "MIGLIOR" θ È PRENDERE QUELLO IN CORRISPONDENZA DEL QUALE LA FUNZIONE $\theta \in \Theta \to L_{\hat{x}}(\theta) = L(\hat{x},\theta)$ (QUI $\hat{x}$ È FISSATO ED È L'OSSERVATO) È MASSIMA. IL PRINCIPIO È CHE "CIÒ CHE È STATO OSSERVATO È CIÒ CHE AVEVA MAGGIOR PROBABILITÀ DI VERIFICARSI". SI SCEGLIE DUNQUE θ COSÌ CHE

$p_\theta(X_1 = x_1, \dots, X_n = x_n)$ SIA MASSIMA IN $x = \hat{x}$. IN EFFETTI NEL CASO DISCRETO ABBIAMO

$$L(x, \theta) = p_\theta(X_1 = x_1, \dots, X_n = x_n)$$

PER APPROSSIMAZIONE SI MASSIMIZZA $L_{\hat{x}}(\theta)$ ANCHE NEL CASO ASS. CONTINUO.

LA FUNZIONE $\theta \to L_{\hat{x}}(\theta)$ È DETTA <u>FUNZIONE DI VERO-SIMIGLIANZA</u> (LIKELIHOOD IN INGLESE).

UNO STIMATORE $H = h(X)$ È DETTO DI <u>MASSIMA VERO-SIMIGLIANZA</u> SE PER OGNI VALORE x DELL'OSSERVAZIONE LA FUNZIONE DI VEROSIMIGLIANZA $\theta \to L_x(\theta)$ RAGGIUNGE IL SUO MASSIMO PER $\theta = h(x)$

<u>NOTA</u>: UNO STIMATORE DI MASSIMA VEROSIMIGLIANZA PUÒ NON ESISTERE (SE IN CORRISPONDENZA DEL VALORE x LA FUNZIONE $L_x(\theta)$ NON HA MASSIMO ASSOLUTO) OPPURE NON ESSERE UNICO (SE $L_x(\theta)$ HA PIÙ PUNTI DI MASSIMO).

COME PROCEDIAMO NELLA PRATICA? SE Θ È UN APERTO E $L_x(\theta)$ È REGOLARE, SI PROCEDE ALLA RICERCA DEGLI

ZERI DI L'_x (SE $\Theta \subset \mathbb{R}$) O PIÙ IN GENERE DI ∇L_x (SE $\Theta \subset \mathbb{R}^m$). SPESSO È PIÙ COMODO* CERCARE I PUNTI DI MASSIMO DI $\theta \to \log L_x(\theta)$: POICHÉ $\log$ È CRESCENTE, I PUNTI DI MASSIMO SONO GLI STESSI.

UNA VOLTA TROVATO $\theta(x)$ TALE CHE $\nabla L_x(\theta(x)) = 0$, OCCORRE VERIFICARE CHE EFFETTIVAMENTE

$$L_x(\theta(x)) = \max\{L_x(\theta) : \theta \in \Theta\}$$

E NEL CASO SI PONE $h(x) = \theta(x)$.

SIA $(X_1, \dots, X_n)$ IL CAMPIONE DA UTILIZZARE E SIA $\ell(x, \theta)$ LA DENSITÀ DI OGNUNA DELLE X_k (SONO EQUIDISTRIBUITE). ESSENDO INDIPENDENTI,

$$L(x, \theta) = \prod_{k=1}^{n} \ell(x_k, \theta)$$

(* QUI SE PASSIAMO AL LOGARITMO CI TROVIAMO UNA SOMMA)

VOLENDO MASSIMIZZARE $\theta \to \log L_x(\theta)$, E SUPPONENDO CHE θ SIA 1-DIMENSIONALE, IMPONIAMO

$$\frac{d}{d\theta} \log L_x(\theta) = \frac{d}{d\theta} \sum_{k=1}^{n} \log \ell(x_k, \theta) = 0$$

TROVANDO

$$\boxed{\sum_{k=1}^{n} \frac{d\ell(x_k, \theta)}{d\theta} \cdot \frac{1}{\ell(x_k, \theta)} = 0}$$

EQUAZIONE DI VEROSIMIGLIANZA

ESEMPIO: SUPPONIAMO DI AVER FATTO n OSSERVAZIONI SU UNA POPOLAZIONE CHE SI ASSUME SEGUIRE UNA LEGGE NORMALE. ABBIAMO QUINDI UN CAMPIONE $(X_1, \dots, X_n)$ CON $X_k \sim \mathcal{N}(\mu, \sigma^2)$. VOGLIAMO STIMARE μ. POICHÉ LE X_k SONO ASSUNTE INDIPENDENTI

$$L_x(\mu) = \underbrace{\prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_k-\mu)^2}{2\sigma^2}\right]}_{\text{DENSITÀ CONGIUNTA}}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\sum_{k=1}^{n} \frac{(x_k-\mu)^2}{2\sigma^2}\right]$$

GUARDANDO L'EQUAZIONE DI VEROSIMIGLIANZA, DOBBIAMO PRIMA CALCOLARE LE DERIVATE

$$\frac{d\,l(x_k,\mu)}{d\mu} = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{x_k-\mu}{\sigma^2} \exp\left[-\frac{(x_k-\mu)^2}{2\sigma^2}\right] = \frac{x_k-\mu}{\sigma^2}\, l(x_k,\mu)$$

E POI IMPORRE $\quad \frac{1}{\sigma^2} \sum_{k=1}^{n} (x_k-\mu) = 0$

NE RICAVIAMO $\left(\sum_{k=1}^{n} x_k\right) - n\mu = 0$, DA CUI $h(x) = \frac{1}{n} \sum_{k=1}^{n} x_k$

E QUINDI LA MEDIA CAMPIONARIA È UNO STIMATORE DI MASSIMA VEROSIMIGLIANZA, CONOSCENDO OPPURE NO LA VARIANZA σ^2.

ESEMPIO Sia $(X_1, \dots, X_n)$ un campione con $X_k \sim \mathcal{P}(\lambda)$ (distribuzione di Poisson). Stimare λ.

Se $x = (x_1, \dots, x_n)$ è uno dei possibili valori osservati (gli x_k devono essere degli interi), allora

$$L_x(\lambda) = \underbrace{\prod_{k=1}^{n} \frac{\lambda^{x_k}}{x_k!} e^{-\lambda}}_{\text{DENSITÀ CONGIUNTA}} = e^{-n\lambda} \frac{\lambda^{\sum_{k=1}^{n} x_k}}{\prod_{k=1}^{n} x_k!}$$

Il suo logaritmo, a meno di una costante additiva che non dipende da λ, è

$$\lambda \to -n\lambda + \left(\sum_{k=1}^{n} x_k\right) \log \lambda$$

Derivando (rispetto λ), un punto critico dovrà risolvere

$$-n + \frac{1}{\lambda}\left(\sum_{k=1}^{n} x_k\right) = 0$$

Ne segue che $h(x) = \frac{1}{n} \sum_{k=1}^{n} x_k$ e quindi la media campionaria è uno stimatore di massima verosimiglianza per λ. Potevamo aspettarcelo: se $X \sim \mathcal{P}(\lambda)$, allora $E[X] = \lambda$.

<u>ESEMPIO</u> SIMILE ALL'ESEMPIO PRECEDENTE, MA QUESTA VOLTA $X_u \sim \mathcal{E}(\lambda)$ (DISTRIBUZIONE ESPONENZIALE).

$$L_x(\lambda) = \prod_{u=1}^{m} \lambda e^{-\lambda x_u} = \lambda^m e^{-\lambda \sum_{u=1}^{m} x_u}$$

$$\log L_x(\lambda) = m \log \lambda - \lambda \sum_{u=1}^{m} x_u$$

DERIVANDO SI OTTIENE L'EQUAZIONE DI VEROSIMIGLIANZA

$$\frac{m}{\lambda} - \sum_{u=1}^{m} x_u = 0$$

DA CUI

$$h(x) = \frac{1}{\frac{1}{m} \sum_{u=1}^{m} x_u}$$

CIOÈ IL RECIPROCO DELLA MEDIA CAMPIONARIA È LO STIMATORE DI MASSIMA VEROSIMIGLIANZA PER λ.

ANCHE IN QUESTO CASO POTEVAMO ASPETTARCELO: SE $X \sim \mathcal{E}(\lambda)$ ALLORA $E[X] = 1/\lambda$.

<u>NOTA</u>: UN RISULTATO IDENTICO VALE SE $X_u \sim \mathrm{Geo}(q)$ (DISTRIBUZIONE ESPONENZIALE). IN EFFETTI ANCHE IN QUESTO CASO $E[X] = 1/q$

NOTA BENE Anche se gli esempi mostrati fanno sperare che lo stimatore di massima verosimiglianza sia corretto, non sempre è così.

ESEMPIO Torniamo al caso gaussiano. Questa volta vogliamo stimare σ.

$$L_x(\sigma) = \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left[-\sum_{n=1}^{M} \frac{(x_n-\mu)^2}{2\sigma^2}\right]$$

$$\log L_x(\sigma) = -\frac{M}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{n=1}^{M}(x_n-\mu)^2$$

$$\frac{d\log L_x(\sigma)}{d\sigma^2} = \frac{-M}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{n=1}^{M}(x_n-\mu)^2$$

Questa derivata si annulla per $\sigma^2 = \frac{1}{M}\sum_{n=1}^{M}(x_n-\mu)^2$

Quindi, se conosciamo μ, la v.a $H = \frac{1}{M}\sum_{n=1}^{M}(X_n-\mu)^2$ è uno stimatore di massima verosimiglianza e, come sappiamo, è corretto.

Se invece entrambe μ e σ sono incognite, allora $\theta = (\mu,\sigma) \in \mathbb{R}^2$ ed il problema diventa vettoriale: dobbiamo considerare $\nabla \log L_x(\theta)$ ed imporre

$$
\begin{cases}
\dfrac{\partial L_x(\theta)}{\partial \mu} = 0 \\[2ex]
\dfrac{\partial L_x(\theta)}{\partial \sigma^2} = 0
\end{cases}
$$

LA PRIMA EQUAZIONE RESTITUISCE $\mu = \frac{1}{n}\sum_{k=1}^{n} x_k$, CHE REINSERITA NELLA SECONDA DA $\sigma^2 = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \frac{1}{n}\sum_{k=1}^{n} x_k\right)^2$

QUESTO CI DICE CHE $H = \frac{1}{n}\sum_{k=1}^{n}\left(X_k - \bar{X}_n\right)$ È LO STIMATORE DI MASSIMA VEROSIMIGLIANZA PER σ, ANCHE NON ESSENDO CORRETTO.

<u>ESEMPIO</u> SIMILE AGLI ESEMPI PRECEDENTI, MA QUESTA VOLTA $X_k \sim \Gamma(\alpha, \lambda)$, α NOTA. POICHÉ $E[X] = \alpha/\lambda$, CI ASPETTIAMO CHE $H = \alpha/\bar{X}$ SIA LO STIMATORE DI MASSIMA VEROSIMIGLIANZA PER λ. COSÌ È, BASTA FARE I CONTI.