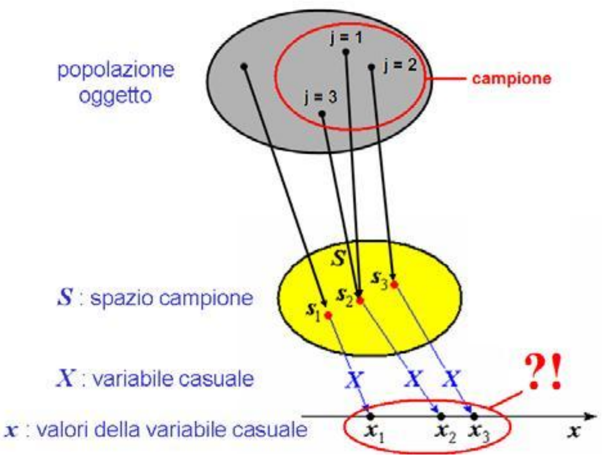
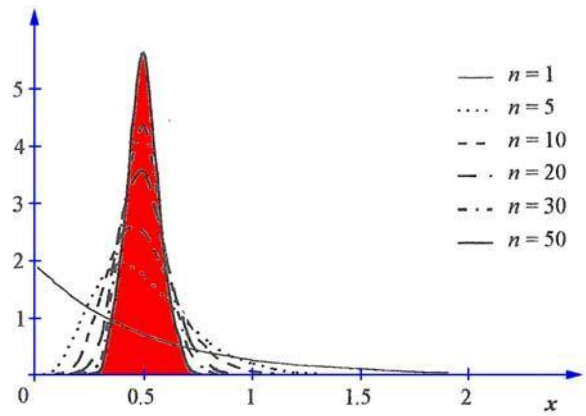
FACOLTA' DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA CIVILE

CORSO DI IDROLOGIA

PROF. PASQUALE VERSACE





SCHEDA DIDATTICA N°5

MOMENTI DELLE VARIABILI CASUALI E STIMA DEI PARAMETRI

A.A. 2012-13

Momenti delle variabili casuali

La distribuzione di probabilità di una variabile casuale X, discreta (o continua), è descritta in modo completo dalla funzione di ripartizione associata, o dalla corrispondente funzione (di densità) di probabilità. Nonostante ciò, spesso si è interessati a conoscere soltanto alcuni aspetti parziali della distribuzione di probabilità di X, specie quando si vogliano confrontare diverse distribuzioni. Sono utili in questo caso valori di sintesi che descrivono aspetti specifici della distribuzione di probabilità riferita a X, ad esempio la posizione o la dispersione sulla retta reale.

Valore atteso o valore medio

Si chiama valore atteso (valor medio o media) di una variabile casuale discreta (o continua) X, la media dei suoi possibili valori, pesati con le relative probabilità (o funzione di densità di probabilità), ovvero:

$$E[X] = \sum x_i P(X = x_i) \qquad \text{per v.c. discrete}$$

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx \qquad \text{per v.c. continue}$$

Usualmente si indica $E[X] = \mu$.

Esempio 1

Tizio e Caio giocano al seguente gioco. Si lancia una moneta, se esce testa Tizio paga a Caio 1 euro, se esce croce è Caio a dover dare a Tizio la stessa somma. Se X è la variabile casuale che descrive il guadagno di Tizio, si vede subito che è una variabile discreta che assume solo due valori: X=-1 nel caso esca testa (ha perso 1 euro) e X=1 nel caso esca croce.
Indicata con p la probabilità che esca testa, si vuole determinare il guadagno atteso di Tizio.
Tale quantità vale:

$$E[X] = (-1)p + (1)(1 - p) = 1 - 2p$$

Quindi, $E[X]$ risulta positivo, nullo o negativo se, rispettivamente, $p < 1/2$, $p = 1/2$ o $p > 1/2$. Pertanto, nel caso in cui la moneta sia regolare (p=1/2) il guadagno atteso di Tizio è nullo .

Esempio 2

Si consideri la variabile casuale esponenziale X. Poichè $f_X(x) = \lambda e^{-\lambda x}$, per $x \geq 0$, e nulla altrove, con $\lambda > 0$ una costante reale, si ottiene che:

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^{+\infty} x\lambda e^{-\lambda x} dx = -\int_0^{+\infty} x \frac{d}{dx} e^{-\lambda x} dx = [-xe^{-\lambda x}]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx = 0 - \frac{1}{\lambda}\left[e^{-\lambda x}\right]_0^{+\infty} = \frac{1}{\lambda}$$

avendo applicato l'integrazione per parti.

Varianza

Molto spesso accade che distribuzioni di probabilità aventi lo stesso valore atteso, differiscano sensibilmente tra loro. Può essere utile allora introdurre indici in grado di esprimere, in forma sintetica, ulteriori caratteristiche della distribuzione di probabilità di X, quale la dispersione dei possibili valori di X intorno al valor medio.

Esempio 3a

Consideriamo ancora il caso, già definito nell'esempio1, del lancio di una moneta. Questa volta però la posta in gioco è costituita da 1000 euro. Anche in questo caso, definita Z la v.c. che descrive il guadagno, questa può assumere due valori, -1000 e 1000, ciascuno con probabilità pari a ½ se la moneta è ben tarata. In queste condizioni $E[Z]$, vale 0.

Questo secondo gioco però è molto più rischioso: infatti, nel primo caso si poteva al massimo perdere 1 euro alla volta, ora è possible perdere 1000 euro! Eppure per una moneta regolare $E[Z]=E[X] = 0$.

La differenza fondamentale tra i due esempi considerati è che mentre X assume valori vicini alla propria media, Z assume valori lontani da $E[Z]$ Pertanto $E[X]$ rappresenta X meglio di quanto non faccia $E[Z]$ per Z.

L'indice più comunemente utilizzato per rappresentare la dispersione di una v.c. rispetto alla sua media è la varianza.

Data una variabile casuale X discreta o continua con valore atteso $E[X]$, si chiama varianza di X, e la si indica con σ^2, o anche con $Var[X]$, la quantità

$$Var[X]=E[(X - E[X])^2]$$

Esempio 3b

Sia X il guadagno che si ha giocando a testa e croce puntando 1 euro e Z quello che si ha puntando 1000 euro. Per una moneta ben tarata si ha $P[X = -1] = P[X = 1] = P[Z = -1000] =P[Z = 1000] = 1/2$ e $E[X] = E[Z] = 0$. Per quanto riguarda la varianza di X si ottiene:

$$Var[X] = E[(X - E[X])^2] = E[X^2] = \sum_{x\in\{-1,1\}} x^2 P(X = x) = (-1)^2 \frac{1}{2} + 1^2 \frac{1}{2} = 1$$

mentre per quella di Z si ha

$$Var[Z] = E[(Z - E[Z])^2] = E[Z^2] = \sum_{z\in\{-10^3,10^3\}} x^2 P(X = x) = (-10^3)^2 \frac{1}{2} + (10^3)^2 \frac{1}{2} = 10^6$$

Come già anticipato $Var[Z]$ è (molto) più grande di $Var[X]$ ad indicare che Z si discosta da $E[Z]$ molto più di quanto non faccia X da $E[X]$.

Il valore atteso e la varianza di una variabile casuale X costituiscono casi particolari dei momenti di una v.c, che saranno di seguito brevemente richiamati.

Momenti della v.c. *X*

Sia X una variabile casuale e sia $R \in N^+$. Si chiama momento di ordine R di X la quantità $\mu_R = E[X^R]$, ossia il valore atteso della variabile casuale trasformata $g(X) = X^R$.

$\mu_R = E[X^R] = \sum x_i^R p_i$ per v.c. discrete $\left(p_i = P(X = x_i)\right)$

$\mu_R = E[X^R] = \int_{-\infty}^{+\infty} x^R f_X(x)dx$ per v.c. continue

La media corrisponde, quindi, al momento primo di $X \Rightarrow \mu_1 = \mu$

Momenti della v.c. Scarto

Tramite il momento primo μ è possibile definire Y, v.c. scarto rispetto alla media:

$$Y=(X - \mu).$$

Il momento di ordine R di Y, variabile casuale trasformata scarto, risulta.

$\mu_R^{'} = E[Y^R] = E[(X - \mu)^R] = \sum (x_i - \mu)^R p_i$ per v.c. discrete

$\mu_R^{'} = E[Y^R] = E[(X - \mu)^R] = \int_{-\infty}^{+\infty} (x - \mu)^R f_X(x)dx$ per v.c. continue

Le quantità definite sono anche indicate come momenti centrali di ordine R di X rispetto a μ. Tali momenti rappresentano una misura della variabilità della distribuzione rispetto alla media. Il più importante, come visto, è il momento che si ottiene per $R=2$, cioè la varianza:

$$\mu_2^{'} = \sigma^2 = E[(X - \mu)^2] = Var[X]$$

Osservazioni

- Il momento del primo ordine della variabile casuale scarto vale sempre 0. Infatti,

$$E[(X - \mu)] = E[X] - E[\mu] = \mu - \mu = 0$$

- Se X è una variabile casuale degenere, ovvero se assume un unico valore con probabilità 1, la varianza è nulla, mentre è tanto più elevata quanto maggiore è la dispersione dei valori di X attorno a μ.
- E' immediato verificare che $Var[X] \geq 0$.
- Si dimostra che è possibile esprimere la varianza rispetto ai momenti di X:

$$Var[X] = E[(X - \mu)^2] = E[X^2 - 2X\mu + \mu^2] = E[X^2] - 2E[X]\mu + \mu^2 =$$

$$= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2$$

ovvero la varianza è uguale al momento secondo ($E[X^2]=\mu_2$) meno il momento primo al quadrato ($\mu^2=\mu_1^2$) della variabile X.

Esempio 4
Si consideri la variabile casuale discreta X che riporta il numero degli esiti testa in tre lanci di una moneta regolare. I valori che X può assumere sono $\{0, 1, 2, 3\}$ e si ha:

$P(X = 0) = P(X = 3) = 1/8$,

$P(X = 1) = P(X = 2) = 3/8$.

E' facile verificare che:

$$\mu_1 = E[X] = 0 + 1\frac{3}{8} + 2\frac{3}{8} + 3\frac{1}{8} = \frac{3}{2}$$

$$\mu_2 = E[X^2] = 0 + 1\frac{3}{8} + 4\frac{3}{8} + 9\frac{1}{8} = 3$$

Quindi, utilizzando la regola per il calcolo della varianza dai momenti di X sopra riportata, si ottiene che $Var[X] = \mu_2 - \mu_1^2 = 3/4$.
Al medesimo risultato si giunge applicando direttamente la definizione di varianza come momento del secondo ordine della variabile scarto.

$$Var[X] = \left(0 - \frac{3}{2}\right)^2 \frac{1}{8} + \left(1 - \frac{3}{2}\right)^2 \frac{3}{8} + \left(2 - \frac{3}{2}\right)^2 \frac{3}{8} + \left(3 - \frac{3}{2}\right)^2 \frac{1}{8} = 3/4.$$

La radice quadrata non negativa della varianza, indicata con σ, è detta *deviazione standard* o *scarto quadratico medio*. Anch'essa costituisce una misura della dispersione di X attorno a μ, spesso preferita alla varianza in quanto è espressa nella stessa unità di misura della variabile casuale X.
Si può, inoltre, definire la quantità σ/μ, chiamata *coefficiente di variazione*. Poichè il coefficiente di variazione non dipende dall'unità di misura con cui viene studiato il fenomeno, può risultare utile per confrontare la dispersione di due o più variabili casuali.

Momenti della v.c. Standardizzata
Data una variabile casuale X, con $\mu = E[X]$ e $\sigma^2 = Var[X]$ è possibile definire la variabile casuale trasformata Z tale che:

$$Z = \frac{X - \mu}{\sigma}.$$

Z è detta variabile casuale standardizzata, ed è caratterizzata dall'essere svincolata dal valor medio e indipendente dalla variabilità misurata dalla varianza. Si può in effetti dimostrare che $E[Z] = 0$ e $Var[Z]=1$.
Di seguito sono riportate l'espressioni dei momenti di ordine R della v.c. standardizzata Z:

$$\overline{\overline{\mu_R}} = E[Z^R] = E\left[\left(\frac{X - \mu}{\sigma}\right)^R\right] = \sum \left(\frac{x_i - \mu}{\sigma}\right)^R p_i \qquad \text{per v.c. discrete}$$

$$\overline{\overline{\mu_R}} = E[Z^R] = E\left[\left(\frac{X-\mu}{\sigma}\right)^R\right] = \int_{-\infty}^{+\infty}\left(\frac{x-\mu}{\sigma}\right)^R f_X(x)dx \qquad \text{per v.c. continue}$$

Tra i momenti della v.c. standardizzata alcuni assumono particolare interesse nel fornire indicazioni sulla forma della distribuzione. In particolare, il coefficiente di asimmetria $\overline{\overline{\mu_3}}$, spesso indicato con il simbolo γ_1, ed il coefficiente di curtosi $\overline{\overline{\mu_4}} - 3$, spesso indicato con il simbolo γ_2.

Osservazioni

- Il coefficiente di asimmetria fornisce indicazioni rispetto alla simmetria della distribuzione rispetto alla media.

$$\gamma_1 = \overline{\overline{\mu_3}} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

  L'indice è indipendente dall'unità di misura della v.c. e può assumere valori negativi, nulli o positivi. E' nullo se la distribuzione è simmetrica rispetto a μ, è negativo se la distribuzione è asimmetrica negativa (coda a sinistra), è positivo se la distribuzione è asimmetrica positiva (coda a destra).

- Il coefficiente di curtosi misura il grado di appiattimento della distribuzione rispetto alla distribuzione normale.

$$\gamma_2 = \overline{\overline{\mu_4}} - 3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$$

  Anche questo indice è indipendente dall'unità di misura e può essere negativo, nullo o positivo. Se γ_2 è nullo, si dice che *X* si distribuisce in modo abbastanza simile ad una normale con stessa media e varianza di *X*; se è maggiore di zero, si dice che la distribuzione è *leptocurtica*, cioè più appuntita della distribuzione normale; se è minore di zero, si dice che la distribuzione è *platicurtica*, cioè più piatta della corrispondente distribuzione normale.

Momenti	Definizione Generale	Formule per v.c. discrete	Formule per v.c. continue
v.c. originaria X	$\mu_R = E[X^R]$ Valore medio: $\mu_1 = \mu = E[X]$	$\mu_R = \sum x_i^R p_i$	$\mu_R = \int_{-\infty}^{+\infty} x^R f_X(x)dx$
v.c. scarto $Y=X-\mu$	$\mu_r^{'} = E[(X-\mu)^r]$ Varianza: $\mu_2^{'} = \sigma^2 = E[(X-\mu)^2]$	$\mu_R^{'} = \sum (x_i - \mu)^R p_i$	$\mu_R^{'} = \int_{-\infty}^{+\infty} (x-\mu)^R f_X(x)dx$
v.c. standardizzata $Z = \dfrac{X-\mu}{\sigma}$	$\overline{\overline{\mu}}_R = E\left[\left(\dfrac{X-\mu}{\sigma}\right)^R\right]$ Asimmetria: $\overline{\overline{\mu_3}} = \gamma_1$ Curtosi: $\overline{\overline{\mu_4}} - 3 = \gamma_2$	$\overline{\overline{\mu_R}} = \sum \left(\dfrac{x_i-\mu}{\sigma}\right)^R p_i$	$\overline{\overline{\mu_R}} = \int_{-\infty}^{+\infty} \left(\dfrac{x-\mu}{\sigma}\right)^R f_X(x)dx$

Altri indici sintetici di posizione

Moda

Data una variabile casuale X, si chiama **moda** della distribuzione di probabilità di X, o più semplicemente moda di X, il valore reale per cui è massima la funzione (di densità) di probabilità, cioè tale che:

$f_X(x_{mo}) \geq f_X(x)$, per ogni x.

E' opportuno osservare che la moda non è necessariamente unica e può anche non esistere.
Se esiste individua i valori più probabili, se X è discreta, o i valori nel cui intorno ricadono gli eventi più probabili, se X è continua. Nel caso in cui si ha un unico massimo, la distribuzione (di densità) di probabilità di X è detta unimodale; se ci sono due o più punti di massimo, si parla di distribuzioni bimodali o multimodali.

Mediana

Data una variabile casuale continua X, si chiama **mediana** della distribuzione di probabilità di X, o più semplicemente mediana di X, e si indica con $x_{0.5}$, il valore per il quale la funzione di ripartizione vale 0.5.

La definizione di mediana può essere interpretata come un caso particolare della definizione più generale di quantile.

Quantili

Sia $\alpha \in (0;1)$ e X una variabile casuale continua. Si chiama quantile (o frattile) α-esimo della distribuzione di probabilità di X, o più semplicemente quantile α-esimo di X, e si indica con x_α, il valore x_α tale che la funzione di ripartizione risulta

$F_X(x_\alpha)=\alpha$.

In questo contesto, x_α è interpretabile come quel valore reale che ripartisce la massa unitaria di probabilità riferita alla variabile casuale X, lasciando una porzione pari ad α alla propria sinistra e pari a $1-\alpha$ alla propria destra. Solitamente α è espresso in termini decimali o percentuali e si parla allora di decili o di percentili. La mediana, quindi, costituisce il 50-esimo percentile o, analogamente, l'α-esimo quantile, con $\alpha = 0.5$.

I quantili x_α, con $\alpha = 1/4; 1/2; 3/4$, sono anche chiamati quartili.

La distanza tra il primo ed il terzo quartile, misurata con la differenza $x_{3/4} - x_{1/4}$, spesso è utilizzata come misura sintetica della dispersione.

MOMENTI DELLE PRINCIPALI DISTRIBUZIONI DI PROBABILITÀ

Distribuzioni discrete

La distribuzione di Bernoulli

Parametri: p $\quad 0 \le p \le 1$

Momenti

Media	p
Varianza	$(1\text{-}p)p$
Coefficiente di variazione	$\frac{\sqrt{p(1-p)}}{p}$
Coefficiente di asimmetria	$\frac{1-2p}{\sqrt{(1-p)p}}$
Kurtosi	$3+\frac{1-6p(1-p)}{\sqrt{(1-p)p}}$

La distribuzione Binomiale

Parametri: p, n $\quad 0 \le p \le 1;\, n \ge 0$

Momenti

Media	np
Varianza	$np(1-p)$
Coefficiente di variazione	$\sqrt{\frac{1-p}{np}}$
Coefficiente di asimmetria	$\frac{1-2p}{\sqrt{np(1-p)}}$
Kurtosi	$3+\frac{1-6p+6p^2}{\sqrt{np(1-p)}}$

La distribuzione Geometrica

Parametri: p $\quad 0 \le p \le 1$

Momenti

Media	$\frac{1}{p}$
Varianza	$\frac{(1-p)}{p^2}$
Coefficiente di variazione	$\sqrt{1-p}$

La distribuzione Binomiale Negativa

Parametri: p, k $\quad 0 \le p \le 1;\, k \ge 1$

Momenti

Media	$\frac{k}{p}$
Varianza	$\frac{k(1-p)}{p^2}$
Coefficiente di variazione	$\sqrt{\frac{1-p}{k}}$

La distribuzione di Poisson

Parametri: ν

Momenti

Media	ν
Varianza	ν
Coefficiente di variazione	$\frac{1}{\sqrt{\nu}}$
Coefficiente di asimmetria	$\frac{1}{\sqrt{\nu}}$
Kurtosi	$3+\frac{1}{\nu}$

La distribuzione Uniforme discreta

Parametri: N

Momenti

Media	$\frac{N+1}{2}$
Varianza	$\frac{(N^2-1)}{12}$

Distribuzioni continue

La distribuzione Uniforme Continua o Rettangolare

Parametri: a, b

Momenti

Media	$\frac{a+b}{2}$
Varianza	$\frac{(b-a)^2}{12}$

La distribuzione Normale

Parametri: μ, σ

Momenti

Media	μ
Varianza	σ^2
Coefficiente di variazione	$\frac{\sqrt{\sigma^2}}{\mu}$
Coefficiente di asimmetria	0
Kurtosi	3

La distribuzione LogNormale

Parametri: μ_y, σ_y

Momenti

Media	$e^{\mu_y+\frac{1}{2}\sigma_y^2}$
Varianza	$e^{2\mu_y+2\sigma_y^2}-e^{\mu_y+\sigma_y^2}$

La distribuzione Esponenziale

Parametri: λ $\lambda>0$

Momenti

Media	$\frac{1}{\lambda}$
Varianza	$\frac{1}{\lambda^2}$
Coefficiente di asimmetria	2
Kurtosi	9

La distribuzione Gamma

Parametri: α, β $\alpha>0$; $\beta>0$

Momenti

Media	$\alpha\beta$
Varianza	$\alpha\beta^2$
Coefficiente di asimmetria	$\frac{2}{\sqrt{\alpha}}$
Kurtosi	$3+6/\alpha$

La distribuzione di Gumbel

Parametri: α, ε   α>0; ε>0

Momenti

Media	$\varepsilon + \frac{0.5772}{\alpha}$
Varianza	$\frac{\pi^2}{6\alpha^2}$
Coefficiente di variazione	$\frac{\pi}{\sqrt{6}(\varepsilon\alpha + 0.5772)}$
Coefficiente di asimmetria	1.14

Stima dei parametri

Metodo dei momenti

Da un punto di vista concettuale il metodo dei momenti è la tecnica più semplice di stima dei parametri di una distribuzione. Non richiede la conoscenza della distribuzione della popolazione di cui si vogliono stimare i parametri, ma solo delle relazioni tra questi ed i momenti della popolazione.
La logica del metodo consiste nell'ipotizzare che i momenti della popolazione coincidano con i corrispondenti momenti campionari, ottenuti dalle osservazioni $x_1, x_2, \ldots, x_n$.

Supponiamo che una v.c. abbia una certa funzione (di densità) di probabilità $f_X(x; \theta_1, \theta_2, \ldots, \theta_p)$ in cui $\theta_1, \theta_2, \ldots, \theta_p$ sono i parametri incogniti da stimare. Se si considerano i primi p momenti della popolazione, cioè:

$\mu_R = \int_{-\infty}^{+\infty} x^R f_X(x; \theta_1, \theta_2, \ldots, \theta_p) dx$ nel caso di v.c. continua o,

$\mu_R = \sum x_i^R p_i$ nel caso di v.c. discreta,

con $R = 1, 2, \ldots, p$, questi sono in generale funzione dei p parametri incogniti:

$$\mu_R = \mu_R(\theta_1, \theta_2, \ldots, \theta_p).$$

Indicando i momenti stimati dal campione come $M_R = \sum_{i=1}^{n} \frac{x_i^R}{n}$ con $R = 1, 2, \ldots, p$, è possibile utilizzare il seguente sistema di p equazioni in p incognite per stimare i parametri $\theta_1, \theta_2, \ldots, \theta_p$:

$$\begin{cases} M_1 = \mu_1(\theta_1, \theta_2, \ldots, \theta_p) \\ \ldots \\ M_p = \mu_p(\theta_1, \theta_2, \ldots, \theta_p) \end{cases}$$

Esempio 5

Sia $x_1, x_2, \ldots, x_n$ un campione casuale estratto da un popolazione normale con media μ e varianza σ^2; questi ultimi coincidono con i parametri θ_1 e θ_2 da stimare con il metodo dei momenti. Ricordando che $\sigma^2 = \mu_2 - \mu_1^2$ e $\mu_1 = \mu$. Le equazioni del metodo dei momenti diventano:

$$\begin{cases} M_1 = \mu_1 = \mu \\ M_2 = \mu_2 = \sigma^2 + \mu_1^2 \end{cases}$$

da cui:

$$\begin{cases} \mu = M_1 = \sum_{i=1}^{n} \frac{x_i}{n} \\ \sigma^2 = M_2 - M_1^2 = \sum_{i=1}^{n} \frac{x_i^2}{n} - \left(\sum_{i=1}^{n} \frac{x_i}{n} \right)^2 \end{cases}$$

La media e la varianza della popolazione risultano quindi coincidenti con la media e la varianza campionaria.

Metodo della Massima verosimiglianza

Sia $x_1, x_2, \ldots, x_n$ il campione di una v.c. X che assumiamo provenire da una distribuzione con funzione di densità $f_X(x; \boldsymbol{\theta})$ nel caso di v.c. continua (per brevità si è indicato con $\boldsymbol{\theta}$ il vettore dei parametri $\theta_1, \theta_2, \ldots, \theta_p$ che la caratterizzano), o distribuzione di probabilità $p(x; \boldsymbol{\theta})$ nel caso di v.c. discreta.

Nell'ipotesi che le osservazioni siano indipendenti, una misura della probabilità di avere ottenuto proprio quel campione da una popolazione con la distribuzione considerata, è fornita dalla seguente funzione:

$$L(\boldsymbol{\theta}; x_1, x_2, \ldots, x_n) = f_X(x_1; \boldsymbol{\theta})\, f_X(x_2; \boldsymbol{\theta}) \ldots f_X(x_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f_X(x_i; \boldsymbol{\theta}) \qquad \text{se } X \text{ è continua}$$

$$L(\boldsymbol{\theta}; x_1, x_2, \ldots, x_n) = p(x_1; \boldsymbol{\theta})\, p(x_2; \boldsymbol{\theta}) \ldots p(x_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} p(x_i; \boldsymbol{\theta}) \qquad \text{se } X \text{ è discreta}$$

che è detta funzione di verosimiglianza

Il metodo della massima verosimiglianza consiste nello scegliere come valori $\hat{\boldsymbol{\theta}}$ dei parametri quelli che massimizzano $L(\boldsymbol{\theta}; x_1, x_2, \ldots, x_n)$.

$$L(\hat{\boldsymbol{\theta}}; x_1, x_2, \ldots, x_n) \geq L(\boldsymbol{\theta}; x_1, x_2, \ldots, x_n)$$

Si osservi che, nel caso di distribuzioni discrete, $L(\boldsymbol{\theta}; x_1, x_2, \ldots, x_n)$ è proprio la probabilità di avere ottenuto il campione $x_1, x_2, \ldots, x_n$. Nel caso di distribuzioni continue, per le quali la probabilità di un particolare insieme finito di valori è comunque nulla, L è approssimativamente proporzionale alla probabilità dell'estrazione di un campione di n elementi, $y_1, y_2, \ldots, y_n$, con $y_i \in [x_i - \varepsilon, x_i + \varepsilon]$, $i=1,2,\ldots,n$ e con ε opportunamente piccolo.

Se si ha un solo parametro, $\boldsymbol{\theta} = \theta_1$, lo stimatore di massima verosimiglianza è soluzione dell'equazione:

$$\frac{dL(\theta_1)}{d\theta_1} = 0 .$$

Se la funzione ha p parametri, $\boldsymbol{\theta} = \theta_1, \theta_2, \ldots, \theta_p$, allora il punto che rende massima la funzione di verosimiglianza è una soluzione delle p equazioni:

$$\frac{\partial L(\theta_1,...,\theta_p)}{\partial \theta_1} = 0$$

$$\frac{\partial L(\theta_1,...,\theta_p)}{\partial \theta_2} = 0$$

$\dots$

$$\frac{\partial L(\theta_1,...,\theta_p)}{\partial \theta_p} = 0$$

Osservazioni

- A differenza del metodo dei momenti, quello della massima verosimiglianza implica la conoscenza della distribuzione della v.c. di cui si vogliono stimare i parametri.
- $L(\boldsymbol{\theta})$ e log $L(\boldsymbol{\theta})$ hanno i loro massimi per lo stesso set di parametri $\boldsymbol{\theta}$**,** ed a volte è più facile trovare il massimo del logaritmo della funzione di verosimiglianza.

Esempio 6

Supponiamo di voler stimare con il metodo della massima verosimiglianza il valore del parametro λ di una distribuzione esponenziale.
La funzione di verosimiglianza risulta:

$$L(\lambda; x_1, x_2, ..., x_n) = \left(\lambda e^{-\lambda x_1}\right)\left(\lambda e^{-\lambda x_2}\right)...\left(\lambda e^{-\lambda x_n}\right) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} = \lambda^n e^{-\lambda n \bar{x}}$$

La derivata di $L(\lambda; x_1, x_2, ..., x_n)$ rispetto a λ è:

$$\frac{dL(\lambda)}{d\lambda} = n\lambda^{n-1} e^{-\lambda n \bar{x}} - \lambda^n n \bar{x} e^{-\lambda n \bar{x}}$$

uguagliando a zero si ottiene:

$$\lambda = \frac{1}{\bar{x}}.$$

STIMA DEI PARAMETRI DELLE PRINCIPALI DISTRIBUZIONI

La distribuzione Normale

Le stime dei parametri μ e σ della distribuzione normale, sia operando con il metodo dei momenti sia operando con il metodo della massima verosimiglianza, si ottengono mediante le relazioni:

$$\hat{\mu} = \bar{x}$$
$$\hat{\sigma} = s$$

dove $\bar{x}$ e s sono rispettivamente la media campionaria e lo scarto quadratico medio campionario.

La distribuzione LogNormale

Le stime dei parametri μ_y e σ_y della distribuzione lognormale, sia operando con il metodo dei momenti sia operando con il metodo della massima verosimiglianza, si ottengono mediante le relazioni:

$$\hat{\mu}_y = \bar{x}_y$$
$$\hat{\sigma}_y = s_y$$

dove $\bar{x}_y$ e s_y sono rispettivamente la media campionaria e lo scarto quadratico medio campionario della variabile Y=log X.

La distribuzione Esponenziale

La stima del parametro λ della distribuzione esponenziale, sia operando con il metodo dei momenti sia operando con il metodo della massima verosimiglianza, si ottiene mediante la relazione:

$$\hat{\lambda} = \frac{1}{\bar{x}}$$

La distribuzione Gamma

1) Metodo dei Momenti

$$\hat{\beta} = \frac{s^2}{\bar{x}}; \quad \hat{\alpha} = \frac{\bar{x}^2}{s^2}$$

2) Metodo della Massima Verosimiglianza
   Il sistema di equazioni che permette di determinare i parametri della distribuzione Gamma è:

$$\begin{cases} \hat{\beta} = \bar{x} / \hat{\alpha} \\ \log(\bar{x}) - \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \log(x_j) = \log(\hat{\alpha}) - \psi(\hat{\alpha}) \end{cases}$$

dove $\psi(\hat{\alpha})$ = funzione digamma, $\psi(\alpha) = \dfrac{d \log[\Gamma(\alpha)]}{d\alpha}$ e n è la numerosità del campione.

La distribuzione di Gumbel

1) Metodo dei Momenti

$$\hat{\alpha} = \frac{\pi}{\sqrt{6}s}; \quad \hat{\varepsilon} = \bar{x} - \frac{0.5772}{\hat{\alpha}}$$

2) Metodo della Massima Verosimiglianza
Il sistema di equazioni che permette di determinare i parametri della distribuzione di Gumbel è:

$$\begin{cases} \dfrac{1}{\hat{\alpha}} = \bar{x} - \dfrac{\sum_{i=1}^{n} x_i e^{-\hat{\alpha} x_i}}{\sum_{i=1}^{n} e^{-\hat{\alpha} x_i}} \\ e^{-\hat{\alpha}\hat{\varepsilon}} = \dfrac{1}{n} \sum_{i=1}^{n} e^{-\hat{\alpha} x_i} \end{cases}$$