

# Probabilità e Statistica Esercitazioni

a.a. 2006/2007

C.d.L.: Ingegneria per l'Ambiente ed il Territorio, Ingegneria Civile, Ingegneria Gestionale, Ingegneria dell'Informazione

C.d.L.S.: Ingegneria Civile

## *Stima puntuale di parametri*

Ines Campa e Marco Longhi

# Esercizi

*Esercizio 1.* Si consideri per  $\vartheta > 0$  la funzione definita da

$$f(x; \vartheta) = \begin{cases} \frac{4}{\vartheta^2} x & \text{se } 0 \leq x \leq \frac{\vartheta}{2} \\ \frac{4}{\vartheta^2} \cdot (\vartheta - x) & \text{se } \frac{\vartheta}{2} < x \leq \vartheta \\ 0 & \text{altrimenti} \end{cases}$$

1. Verificare che per ogni  $\vartheta > 0$ ,  $f(\cdot; \vartheta)$  rappresenta una funzione di densità di probabilità.
2. Sia  $X_1, X_2, \dots, X_n$  un campione casuale estratto dalla popolazione di densità  $f(\cdot, \vartheta)$ , stabilire se  $T = \frac{2}{n} \sum_{i=1}^n X_i$  è uno stimatore corretto di  $\vartheta$ .

*Risoluzione.* Disegniamo il grafico della funzione di densità di probabilità data

## 1. Risulta

$$\int_{-\infty}^{+\infty} f(x; \vartheta) dx = \frac{1}{2} \vartheta \frac{2}{\vartheta} = 1$$

quindi  $f(\cdot; \vartheta)$  rappresenta una funzione di densità di probabilità per ogni  $\vartheta > 0$

2.  $T$  è uno stimatore **corretto** o **non distorto** di  $\vartheta$  se  $E[T] = \vartheta$ .

Nel nostro caso  $E[X] = \frac{1}{2} \vartheta$  e

$$\begin{aligned} E[T] &= E\left[\frac{2}{n} \sum_{i=1}^n X_i\right] = \frac{2}{n} \sum_{i=1}^n E[X_i] = \\ &= \frac{2}{n} \cdot nE[X_i] = \frac{2}{n} n \frac{1}{2} \vartheta = \vartheta \implies \text{lo stimatore è corretto} \end{aligned}$$

Esercizio 2. Sia  $X$  la variabile casuale di Poisson di parametro  $\lambda$ ,

1. calcolare lo stimatore di massima verosimiglianza di  $\lambda$ ;
2. verificare la correttezza dello stimatore trovato;
3. calcolare l'errore quadratico medio;
4. determinare uno stimatore con il metodo dei momenti;
5. supponendo poi di avere i seguenti dati campionati

2, 3, 2, 4, 3, 5, 4, 2, 3, 2, 3, 4, 2, 3, 4, 3, 2, 3, 4, 5, 4, 3, 3, 4

calcolare il valore stimato di  $\lambda$ .

*Risoluzione.*  $X$  ha funzione di densità di probabilità

$$f_X(x; \lambda) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{se } x \in N \\ 0 & \text{altrove} \end{cases}$$

La funzione di verosimiglianza è

$$L(x_1, \dots, x_n, \lambda) = \frac{\lambda^{x_1}}{x_1!} e^{-\lambda} \cdot \frac{\lambda^{x_2}}{x_2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{x_n}}{x_n!} e^{-\lambda} = \frac{\lambda^{x_1+x_2+\dots+x_n}}{x_1!x_2!\dots x_n!} e^{-n\lambda} =$$
$$= \frac{1}{x_1!x_2!\dots x_n!} \lambda^{\sum_{i=1}^n x_i} \cdot e^{-n\lambda}$$

con  $x_i \in N, i = 1, \dots, n$ . Il valore di  $\lambda$  :  $\hat{\lambda}$  che massimizza  $L$  è detto **stima di massima verosimiglianza** e  $\hat{\Lambda}(X_1, X_2, \dots, X_n)$  è lo **stimatore di massima verosimiglianza**. Calcoliamo

$$\frac{d}{d\lambda} L(x_1, \dots, x_n, \lambda) = \frac{e^{-n\lambda}}{x_1!x_2!\dots x_n!} \left[ \sum_{i=1}^n x_i \lambda^{\sum_{i=1}^n x_i - 1} - n\lambda^{\sum_{i=1}^n x_i} \right]$$

Cerchiamo gli zeri della derivata prima di  $L(x_1, \dots, x_n, \lambda)$  rispetto a  $\lambda$

$$\sum_{i=1}^n x_i \lambda^{\sum_{i=1}^n x_i - 1} - n \lambda^{\sum_{i=1}^n x_i} = 0$$

$$\frac{1}{\lambda} \sum_{i=1}^n x_i \lambda^{\sum_{i=1}^n x_i} - n \lambda^{\sum_{i=1}^n x_i} = 0$$

$$\frac{1}{\lambda} \sum_{i=1}^n x_i - n = 0 \implies \hat{\lambda}(x_1, x_2, \dots, x_n) = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}_n$$

Si verifica che  $\hat{\lambda}$  è un **punto di massimo**, anche se in modo piuttosto faticoso.

Quindi

$$\hat{\lambda}(x_1, x_2, \dots, x_n) = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}_n$$

è la stima di massima verosimiglianza di  $\lambda$  e la media campionaria lo stimatore di massima verosimiglianza. Allo stesso risultato si perviene usando la **log-verosimiglianza**

$$\ln [L(x_1, \dots, x_n, \lambda)] = -n\lambda + \sum_{i=1}^n x_i \ln \lambda - \sum_{i=1}^n \ln(x_i!)$$

Calcoliamo

$$\frac{d}{d\lambda} \ln [L(x_1, \dots, x_n, \lambda)] = -n + \frac{\sum_{i=1}^n x_i}{\lambda}$$

Poniamo uguale a zero la derivata prima

$$-n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0 \implies \hat{\lambda}(x_1, x_2, \dots, x_n) = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}_n$$

Lo stesso risultato di prima, ma si tratta di un punto di massimo?

$$\left. \frac{d^2}{d\lambda^2} \ln [L(x_1, \dots, x_n, \lambda)] \right|_{\lambda = \hat{\lambda}} = -\frac{\sum_{i=1}^n x_i}{\hat{\lambda}^2} < 0 \implies \hat{\lambda} \text{ PUNTO DI MAX}$$

Per determinare se lo stimatore è non distorto calcoliamo

$$E \left[ \hat{\Lambda}(X_1, X_2, \dots, X_n) \right] = E \left[ \bar{X}_n \right] = \frac{1}{n} E \left[ \sum_{i=1}^n X_i \right] = \frac{1}{n} n E[X]$$

Essendo  $X$  v.c. di Poisson di parametro  $\lambda$ , si ha  $E[X] = \lambda$ , quindi

$$E \left[ \hat{\Lambda}(X_1, X_2, \dots, X_n) \right] = \lambda \implies \hat{\Lambda} \text{ stimatore non distorto}$$



Chiamiamo errore quadratico medio di uno stimatore  $T = t(X_1, \dots, X_n)$ , la funzione di  $\vartheta$  data da:

$$\text{MSE}[T](\vartheta) := E[(T - \tau(\vartheta))^2]$$

con  $\tau(\vartheta)$  funzione del parametro da stimare. Ma

$$\text{MSE}[T](\vartheta) := \text{var}[T] + [D[T](\vartheta)]^2, \quad D[T](\vartheta) := \tau(\vartheta) - E[T]$$

Nel nostro caso  $E[\hat{\Lambda}] = \lambda \implies D[\hat{\Lambda}] = 0$ , quindi

$$\text{MSE}[\hat{\Lambda}] = \text{var}[\hat{\Lambda}] = \text{var}[\bar{X}_n] \underbrace{=}_{X_i \text{ ind.}} \frac{1}{n^2} \sum_{i=1}^n \text{var}[X_i] = \frac{1}{n^2} \cdot n\lambda = \frac{\lambda}{n}$$

Determiniamo, ora, uno stimatore **con il metodo dei momenti**. Ricordiamo che con

$$M'_r = \frac{1}{n} \sum_{i=1}^n X_i^r, \quad \mu'_r = E[X^r]$$

indicano rispettivamente il momento campionario assoluto di ordine  $r$  e il momento di ordine  $r$  della variabile casuale  $X$ . Il metodo dei momenti consiste nel risolvere il sistema nelle incognite  $\vartheta_1, \dots, \vartheta_k$  di  $k$ , numero dei parametri incogniti, equazioni

$$\left\{ \begin{array}{l} \mu'_1 = M'_1 \\ \mu'_2 = M'_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu'_k = M'_k \end{array} \right.$$

Nel nostro caso abbiamo solo un parametro da stimare,  $\lambda$ , quindi calcoliamo

$$\mu'_1 = E[X] = \lambda, \text{ e } M'_1 = \frac{1}{n} \sum_{i=1}^n X_i$$

Posto  $\mu'_1 = M'_1$ , risulta  $\bar{\Lambda} = \frac{1}{n} \sum_{i=1}^n X_i$ .

Esercizio 3. Sia  $X_1, X_2 \dots X_n$  un campione casuale di ampiezza  $n$  estratto da una popolazione con distribuzione uniforme nell'intervallo  $[a - b, a + b]$ . Determinare gli stimatori di  $a$  e  $b$  con il metodo dei momenti.

*Risoluzione.* L'intervallo ha ampiezza  $a + b - (a - b) = 2b$ , ne segue che la funzione di densità di probabilità della v.c.  $X$  è

$$f(x; a, b) = \begin{cases} \frac{1}{2b} & \text{se } a - b \leq x \leq a + b \\ 0 & \text{altrove} \end{cases}$$

Per determinare gli stimatori di  $a, b$  con il metodo dei momenti, bisogna risolvere il sistema

$$\begin{cases} \mu'_1 = M'_1 \\ \mu'_2 = M'_2 \end{cases} \implies \begin{cases} \mu'_1 = \bar{X}_n \\ \mu'_2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases}$$

Calcoliamo  $\mu'_1 = \frac{a+b+a-b}{2} = a$  e

$$\mu'_2 = \text{var}[X] + (\mu'_1)^2 = \frac{[(a+b) - (a-b)]^2}{12} + a^2 = \frac{b^2}{3} + a^2$$

Sostituendo

$$\begin{cases} a = \bar{X}_n \\ \frac{b^2}{3} + a^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases}$$

ed eseguendo alcuni passaggi algebrici, ne segue che gli stimatori cercati sono

$$\hat{a} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{e} \quad \hat{b} = \sqrt{\frac{3}{n} \sum_{i=1}^n X_i^2 - \frac{3}{n^2} \left( \sum_{i=1}^n X_i \right)^2}$$

Esercizio 4. Sia data la variabile casuale unidimensionale  $X$  con funzione di densità di probabilità data da:

$$f(x; \vartheta) = \begin{cases} \frac{1}{18\vartheta^4} \cdot e^{-\frac{\sqrt{x}}{3\vartheta^2}} & \text{se } x \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

con  $\vartheta > 0$ .

1. Trovare uno stimatore di  $\vartheta > 0$  con il metodo di massima verosimiglianza.
2. Calcolare la densità di probabilità della variabile casuale  $Y = \sqrt{X}$ .

*Risoluzione.* Costruiamo la funzione di massima verosimiglianza

$$\begin{aligned} L(x_1, \dots, x_n, \vartheta) &= \frac{1}{18\vartheta^4} e^{-\frac{\sqrt{x_1}}{3\vartheta^2}} \cdot \frac{1}{18\vartheta^4} e^{-\frac{\sqrt{x_2}}{3\vartheta^2}} \cdot \dots \cdot \frac{1}{18\vartheta^4} e^{-\frac{\sqrt{x_n}}{3\vartheta^2}} = \\ &= \frac{1}{(18\vartheta^4)^n} e^{-\frac{(\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n})}{3\vartheta^2}} \end{aligned}$$

con  $x_i \geq 0, i = 1, \dots, n$ .

Cerchiamo  $\hat{\vartheta}$  che massimizza  $\ln L$  tra le soluzioni di  $\frac{d}{d\vartheta} \ln L = 0$ . Calcoliamo

$$\begin{aligned}\ln L &= -n \ln 18\vartheta^4 + \ln e^{-\frac{(\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n})}{3\vartheta^2}} \\ &= -n \ln 18 - 4n \ln \vartheta - \frac{\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n}}{3\vartheta^2}\end{aligned}$$

si ha

$$\frac{d}{d\vartheta} \ln L = -\frac{4n}{\vartheta} + \frac{2}{3\vartheta^3} (\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n}) = 0$$

Essendo  $\vartheta > 0$  risulta

$$\hat{\vartheta} = \sqrt{\frac{\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n}}{6n}}$$

Vediamo se si tratta di un punto di massimo, calcoliamo

$$\frac{d^2}{d\vartheta^2} \ln L = +\frac{4n}{\vartheta^2} - \frac{2}{\vartheta^4} (\sqrt{x_1} + \sqrt{x_2} + \dots + \sqrt{x_n})$$

ne segue

$$\frac{d^2}{d\vartheta^2} \ln L \Big|_{\vartheta = \hat{\vartheta}} = \frac{2}{\hat{\vartheta}^2} \left( 2n - \frac{\hat{\vartheta}^2 \cdot 6n}{\hat{\vartheta}^2} \right) = -\frac{8n}{\hat{\vartheta}^2} < 0 \implies \hat{\vartheta} \text{ PUNTO DI MAX}$$

Quindi

$$\hat{\Theta} = \sqrt{\frac{\sum_{i=1}^n \sqrt{x_i}}{6n}}$$

è uno stimatore di  $\vartheta$  di massima verosimiglianza.

Determiniamo la densità di probabilità della variabile casuale  $Y = \sqrt{X}$ . Si ha  $F_Y(y) = P[Y \leq y] = P[\sqrt{X} \leq y]$ . Se  $y < 0$  si ha  $[\sqrt{x} \leq y] = \emptyset$ .

Se  $y \geq 0$ , si ha

$$F_Y(y) = P[X \leq y^2] = \int_0^{y^2} \frac{1}{18\vartheta^4} e^{-\frac{\sqrt{x}}{3\vartheta^2}} dx$$

Posto  $t = \frac{\sqrt{x}}{3\vartheta^2}$ , risulta  $dt = \frac{1}{6\vartheta^2\sqrt{x}} dx \implies 18\vartheta^4 t dt = dx$ . Quindi

$$\int_0^{y^2} \frac{1}{18\vartheta^4} e^{-\frac{\sqrt{x}}{3\vartheta^2}} dx = \int_0^{\frac{y}{3\vartheta^2}} te^{-t} dt = 1 - e^{-\frac{y}{3\vartheta^2}} \left(1 + \frac{y}{3\vartheta^2}\right)$$



Ne segue

$$F_Y(y) \begin{cases} 1 - e^{-\frac{y}{3\vartheta^2}} \left(1 + \frac{y}{3\vartheta^2}\right) & \text{se } y \geq 0 \\ 0 & \text{se } y < 0 \end{cases}$$

derivando si ottiene

$$f_Y(y) \begin{cases} \frac{y}{9\vartheta^4} e^{-\frac{y}{3\vartheta^2}} & \text{se } y \geq 0 \\ 0 & \text{se } y < 0 \end{cases}$$

Esercizio 5. Si supponga che  $X_1, X_2, X_3$  sia un campione casuale di ampiezza 3 estratto da una distribuzione esponenziale di media  $\lambda$ . Si considerino i seguenti stimatori del parametro  $\lambda$ :

$$\Lambda_1 = X_1, \Lambda_2 = \frac{X_1 + X_2}{2}, \Lambda_3 = \frac{X_1 + 2X_2}{3}, \Lambda_4 = \bar{X}_3$$

1. Indicare quali sono gli stimatori non distorti di  $\lambda$ .
2. Individuare tra gli stimatori non distorti quello con MSE minimo.

*Risoluzione.* Sia  $X$  una v.c. esponenziale di media  $\lambda$ , si ha

$$f(x; \lambda) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{se } x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

con  $E[X] = \lambda$  e  $\text{var}[X] = \lambda^2$ .

Calcoliamo il valore atteso degli stimatori dati

$$E[\Lambda_1] = E[X_1] = \lambda$$

$$E[\Lambda_2] = E\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{2}(E[X_1] + E[X_2]) = \frac{1}{2} \cdot 2\lambda = \lambda$$

$$E[\Lambda_3] = E\left[\frac{X_1 + 2X_2}{3}\right] = \frac{1}{3}(E[X_1] + 2E[X_2]) = \frac{1}{3} \cdot (\lambda + 2\lambda) = \lambda$$

$$E[\Lambda_4] = E[\bar{X}_3] = \lambda$$

Tutti gli stimatori dati sono non distorti, per individuare il preferibile, calcoliamo la varianza di ciascuno:

$$\text{var}[\Lambda_1] = \text{var}[X_1] = \lambda^2$$

$$\begin{aligned}\text{var}[\Lambda_2] &= \text{var}\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{4}(\text{var}[X_1] + \text{var}[X_2] + 2\text{cov}[X_1, X_2]) = \\ &= \frac{1}{4} \cdot 2\lambda^2 = \frac{1}{2}\lambda^2 \text{ (} X_1 \text{ e } X_2 \text{ sono indipendenti)}\end{aligned}$$

$$\text{var}[\Lambda_3] = \text{var}\left[\frac{X_1 + 2X_2}{3}\right] = \frac{1}{9}(\text{var}[X_1] + 4\text{var}[X_2]) = \frac{1}{9} \cdot (\lambda^2 + 4\lambda^2) = \frac{5}{9}\lambda^2$$

$$\text{var} [\Lambda_4] = \text{var} [\bar{X}_3] = \frac{1}{3} \lambda^2$$

Lo stimatore non distorto con varianza minima è la media campionaria.

*Esercizio 6.* Il numero di interruzioni (crash) di un personal computer segue una distribuzione di Poisson di parametro  $\lambda$ . Sia  $X_1, X_2, \dots, X_n$  un campione casuale di ampiezza  $n$  estratto dalla distribuzione di Poisson data. Se il costo di riparazione è dato da

$$Y_n = 3\bar{X}_n + \bar{X}_n^2$$

dove  $\bar{X}_n$  è la media campionaria su  $n$  crash, calcolare  $E[Y_n]$ .

*Risoluzione.* Sia  $X$  una v.c. di Poisson di parametro  $\lambda$ , si ha

$$f(x; \lambda) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{se } x = 0, 1, 2, 3, \dots \\ 0 & \text{altrove} \end{cases}$$

con  $E[X] = \lambda$  e  $\text{var}[X] = \lambda$ . Calcoliamo il valore atteso dello stimatore dato

$$E[Y_n] = E[3\bar{X}_n + \bar{X}_n^2] = 3E[\bar{X}_n] + E[\bar{X}_n^2] = 3\lambda + E[\bar{X}_n^2]$$

Ma

$$E \left[ \overline{X}_n^2 \right] - E \left[ \overline{X}_n \right]^2 = \text{var} \left[ \overline{X}_n \right], \quad \text{var} \left[ \overline{X}_n \right] = \frac{\text{var} [X]}{n} = \frac{\lambda}{n},$$

quindi

$$E [Y_n] = 3\lambda + \frac{\lambda}{n} + \lambda^2$$

e passando al limite per  $n \longrightarrow \infty$  si ha

$$E [Y_\infty] = 3\lambda + \lambda^2$$

*Esercizio 7.* Per un corso la prova d'accertamento del profitto è costituita da una batteria di  $\overline{30}$  quesiti con risposta dicotomica. Si può ipotizzare che chi risponde a caso ad un quesito abbia probabilità  $\frac{1}{2}$  di indovinare la risposta esatta. Supponiamo che ogni risposta esatta sia valutata  $\frac{1}{30}$ , si dica qual è per una persona perfettamente ignorante:

1. la probabilità di ottenere la sufficienza in una prova d'esame;
2. il numero medio di prove necessarie per superare l'esame.

Sia  $X$  la variabile casuale che conteggia il numero di risposte esatte in una prova. È ragionevole pensare che  $X = S + I$  dove  $S$  indica la variabile casuale che conteggia il numero dei quesiti di cui il candidato sa la risposta esatta, mentre  $I$  è la variabile casuale che conteggia il numero di risposte indovinate. Sia  $S$  la v.c. binomiale di parametri  $n = 30$  e  $p$  e sia  $I/S = s$  la v.c. binomiale di parametri  $n = 30 - s$  e  $p = \frac{1}{2}$ .

- Si determini la funzione di densità di probabilità congiunta di  $S$  e  $I$ ,  $p_{S,I}(s, i)$ .
- Si mostri che  $p_X(x) = \sum_{s=0}^x p_{S,I}(s, x - s)$  con  $x = 0, \dots, 30$  e dunque  $X$  è una variabile casuale binomiale di parametri  $n = 30$  e  $\frac{1+p}{2}$ .
- Si determini uno stimatore non distorto di  $p$ .

Risoluzione. Per ipotesi  $X$  è la v.c. che conteggia il numero di risposte esatte, ha distribuzione  $Bi(30, \frac{1}{2})$  con

$$p_X(x) = \binom{30}{x} p^x (1-p)^{30-x} \quad x = 0, \dots, 30$$

Nel caso ricerchiamo la probabilità di ottenere appena la sufficienza si ha

$$\bar{p} = p_X(18) = \binom{30}{18} \left(\frac{1}{2}\right)^{18} \left(\frac{1}{2}\right)^{12} \approx 0.08055$$

Se invece vogliamo almeno la sufficienza, dobbiamo calcolare  $P[X \geq 18]$ .  
Dato che

$$\begin{cases} np = 30 \cdot \frac{1}{2} = 15 \geq 5 \\ nq = 30 \cdot \frac{1}{2} = 15 \geq 5 \end{cases} \implies X \text{ ha distribuzione appr. } N\left(15, \frac{15}{2}\right)$$

Quindi  $P[X \geq 18] = P[X \geq 17.5] \approx 0.18066 = \tilde{p}$  con l'interpolazione, facendo invece "la media aritmetica"  $\tilde{p} = 0.1801$ .

$T$  denota il numero di prove di 30 domande ciascuna. Quando si avrà il superamento della prova con esattamente la sufficienza  $T$  è una v.c. di Pascal di parametro  $\bar{p} = 0.08055$ , quindi  $E [T] = \frac{1}{\bar{p}} \approx 12.41465$ , per passare la prova con esattamente la sufficienza bisogna fare mediamente l'esame dalle 12 alle 13 volte.

Se invece si avrà il superamento della prova con almeno la sufficienza  $T$  è una v.c. di Pascal di parametro

$$\left\{ \begin{array}{l} \tilde{p} = 0.18066 \implies E [T] \approx 5.53526 \\ \tilde{p} = 0.1801 \implies E [T] \approx 5.55247 \end{array} \right.$$



Sia  $0 \leq s \leq 30$  e  $0 \leq i \leq 30 - s$  allora

$$\begin{aligned} p_{S,I}(s, i) &= \binom{30}{s} p^s (1-p)^{30-s} \cdot \binom{30-s}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{30-s-i} = \\ &= \binom{30}{s} \cdot \binom{30-s}{i} p^s \left[\frac{1}{2}(1-p)\right]^{30-s} \end{aligned}$$

Scelto  $0 \leq s \leq 30$  e  $0 \leq i \leq 30 - s$  risulta

$$\begin{aligned} p_X(x) &= \sum_{(s,i): s+i=x} p_{S,I}(s, i) = \sum_{(s,i): i=x-s} p_{S,I}(s, x-s) = \\ &= \sum_{s=0}^x \binom{30}{s} \cdot \binom{x}{s} p^s \left[\frac{1-p}{2}\right]^{30-s} = \end{aligned}$$

$$- = \binom{30}{s} \cdot \left(\frac{1-p}{2}\right)^{30-x} \sum_{s=0}^x \binom{x}{s} p^s \left(\frac{1-p}{2}\right)^{x-s}.$$

Ricordando il binomio di Newton

$$\sum_{s=0}^x \binom{x}{s} p^s \left(\frac{1-p}{2}\right)^{x-s} = \left(\frac{1+p}{2}\right)^x$$

si ha

$$p_X(x) = \binom{30}{x} \left(\frac{1-p}{2}\right)^{(30-x)} \left(\frac{1+p}{2}\right)^x,$$

con  $x = 0, 1, \dots, 30$ . Pertanto  $X$  è una binomiale di parametri 30 e  $\frac{1+p}{2}$ , di media

$$E[X] = 30 \cdot \frac{1+p}{2} = 15(1+p).$$

Determiniamo, inoltre, uno stimatore di  $p$  con il metodo dei momenti:

$$\bar{X}_n = 15(1 + p) \implies \hat{p} = \frac{\bar{X}_n}{15} - 1 \quad \text{stimatore di } p .$$

Calcoliamo  $E[\hat{p}]$

$$E[\hat{p}] = \frac{1}{15} \cdot 15(1 + p) - 1 = p,$$

quindi  $\hat{p}$  è corretto.

Esercizio 8.  $X_1, X_2 \dots X_n$  un campione casuale proveniente da una distribuzione rettangolare nell'intervallo  $[0, 2\vartheta]$  con  $\vartheta > 0$ .

1. Si determini lo stimatore di  $\vartheta$  con il metodo dei momenti.
2. Si verifichi se tale stimatore è consistente.

*Risoluzione.* L'intervallo ha ampiezza  $2\vartheta - 0 = 2\vartheta$ , ne segue che la funzione di densità di probabilità della v.c. rettangolare  $X$  è

$$f(x; \vartheta) = \begin{cases} \frac{1}{2\vartheta} & \text{se } 0 \leq x \leq 2\vartheta \\ 0 & \text{altrove} \end{cases}$$

Per determinare lo stimatore di  $\vartheta$  con il metodo dei momenti, bisogna risolvere l'equazione

$$\mu'_1 = M'_1 \implies \mu'_1 = \bar{X}_n,$$

rispetto a  $\vartheta$ .

Calcoliamo  $\mu'_1 = \frac{0+2\vartheta}{2} = \vartheta$ , sostituendo nell'equazione precedente risulta  $\bar{X}_n$  uno stimatore di  $\vartheta$ . Essendo  $E[\bar{X}_n] = \vartheta$ , lo stimatore è non distorto.

Calcoliamo

$$\text{var} [\bar{X}_n] \underbrace{=}_{X_i \text{ indep.}} \frac{1}{n^2} \sum_{i=1}^n \text{var} [X_i] = \frac{1}{n^2} \cdot n \cdot \frac{\vartheta^2}{3} = \frac{\vartheta^2}{3n}.$$

Ne segue

$$\lim_{n \rightarrow +\infty} E \left[ (\bar{X}_n - \vartheta)^2 \right] = \lim_{n \rightarrow +\infty} \frac{\vartheta^2}{3n} = 0,$$

quindi lo stimatore è consistente.

*Esercizio 9* (Tema d'esame del 10/01/2006).

Sia  $X_1, X_2 \dots X_n$  un campione casuale di ampiezza  $n$  estratto da una popolazione con densità di probabilità

$$f_X(x, \vartheta) = \begin{cases} \frac{2}{\vartheta} \left(1 - \frac{x}{\vartheta}\right) & \text{se } 0 \leq x \leq \vartheta \\ 0 & \text{altrove} \end{cases}$$

con  $\vartheta > 0$ .

1. Determinare uno stimatore  $T$  di  $\vartheta$  con il metodo dei momenti.
2. Stabilire se  $T$  è distorto e calcolarne l'errore quadratico medio.

*Risoluzione.* Individuiamo lo stimatore  $T$  del parametro  $\vartheta$ , risolvendo rispetto a  $\vartheta$

$$\bar{X}_n = E[X].$$

Calcoliamo

$$E[X] = \int_0^{\vartheta} \frac{2}{\vartheta} x \left(1 - \frac{x}{\vartheta}\right) dx = \frac{1}{3}\vartheta$$

quindi

$$\bar{X}_n = \frac{1}{3}\vartheta \implies T = 3\bar{X}_n.$$

Lo stimatore è non distorto, infatti,  $E[3\bar{X}_n] = 3E[\bar{X}_n] = 3\frac{\vartheta}{3} = \vartheta$

Calcoliamo, infine, l'errore quadratico medio

$$\text{MSE}[T] = \text{var}[T] = \text{var}[3\bar{X}_n] = 9\frac{\text{var}[X]}{n}.$$

$$E[X^2] = \int_0^{\vartheta} \frac{2}{\vartheta} x^2 \left(1 - \frac{x}{\vartheta}\right) dx = \frac{\vartheta^2}{6} \implies \text{var}[X] = \frac{\vartheta^2}{6} - \frac{\vartheta^2}{9} = \frac{\vartheta^2}{18}.$$

Ne segue  $\text{MSE}[T] = \frac{\vartheta^2}{2n}$ .

Esercizio 10 (Tema d'esame del 12/07/2005).

Sia  $X_1, X_2, \dots, X_n$  un campione casuale di ampiezza  $n$  estratto da una popolazione con densità di probabilità

$$f_X(x, \vartheta) = \begin{cases} 5^{-\vartheta} \vartheta x^{\vartheta-1} & \text{se } 0 < x < 5 \\ 0 & \text{altrove} \end{cases}$$

con  $\vartheta > 0$ . Determinare lo stimatore di massima verosimiglianza  $\hat{\Theta}$  di  $\vartheta$ .

*Risoluzione.* Calcoliamo la funzione di massima verosimiglianza

$$L(x_1, x_2, \dots, x_n, \vartheta) = \prod_{i=1}^n f(x_i, \vartheta) = 5^{-n\vartheta} \vartheta^n \prod_{i=1}^n x_i^{\vartheta-1}$$

con  $0 < x_i < 5$ ,  $i = 1, 2, \dots, n$ . Quindi

$$\ln L = -n\vartheta \ln 5 + n \ln \vartheta + (\vartheta - 1) \sum_{i=1}^n \ln x_i \quad \text{e}$$



$$\frac{d}{d\vartheta} \ln L = -n \ln 5 + \frac{n}{\vartheta} + \sum_{i=1}^n \ln x_i$$

Poniamo

$$\frac{d}{d\vartheta} \ln L = 0 \implies \hat{\vartheta} = \frac{n}{n \ln 5 - \sum_{i=1}^n \ln x_i}$$

Inoltre,

$$\left. \frac{d^2}{d\vartheta^2} \ln L \right|_{\vartheta = \hat{\vartheta}} = -\frac{n}{\hat{\vartheta}^2} < 0$$

quindi

$$\hat{\Theta} = \frac{n}{n \ln 5 - \sum_{i=1}^n \ln x_i}$$

Esercizio 11 (Tema d'esame del 28/06/2005).

Sia  $X_1, \dots, X_n$  un campione casuale, di dimensione  $n$ , estratto da una distribuzione rettangolare uniforme sull'intervallo  $[a, 2a]$ .

1. Determinare uno stimatore  $T_1$  di  $a$  con il metodo dei momenti. Verificare se lo stimatore  $T_1$  è distorto e calcolarne l'errore quadratico medio  $MSE[T_1]$ .
2. Considerato poi lo stimatore  $T_2 = \frac{1}{2}X_1 + \frac{1}{6}X_2$ , verificare se  $T_2$  è distorto e calcolarne l'errore quadratico medio  $MSE[T_2]$ .
3. Supposto  $n = 3$ , quale dei due stimatori  $T_1$  e  $T_2$  di  $a$  è preferibile (giustificare la risposta)?

$$T_1 = \frac{2}{3}\bar{X}_n$$

$T_1$  non distorto

[Risposta  $MSE[T_1] = \frac{a^2}{27n}$  ]

$T_2$  non distorto

$$MSE[T_2] = \frac{5a^2}{216}$$

$T_1$  preferibile

Esercizio 12 (Tema d'esame del 11/04/2006).

Determinare lo stimatore di massima verosimiglianza del parametro  $\lambda$  per un campione estratto dalla densità

$$f(x) = \begin{cases} \lambda^2(x-1)e^{-\lambda(x-1)} & \text{se } x > 1, \\ 0 & \text{altrove.} \end{cases}$$

$$\left[ \text{Risposta } T = \frac{2}{\bar{X}_n - 1} \right]$$

Esercizio 13 (Tema d'esame del 04/07/2006).

Sia  $X_1, \dots, X_8$  un campione aleatorio, di dimensione 8, estratto da una distribuzione rettangolare uniforme sull'intervallo  $[-1, b]$ , con  $b > -1$ .

Si chiede:

1. determinare uno stimatore  $T_1$  di  $b$  con il metodo dei momenti;
2. determinare se lo stimatore  $T_1$  sia distorto;
3. calcolare l'errore quadratico medio  $MSE[T_1]$ ;
4. considerato poi lo stimatore  $T_2 = 4\bar{X}_8 - X_3 - X_5 + 1$ , calcolarne l'errore quadratico medio  $MSE[T_2]$ ;
5. determinare quale dei due stimatori  $T_1$  e  $T_2$  di  $b$  sia preferibile, giustificando la risposta.

$$T_1 = 2\bar{X}_n + 1$$

$T_1$  non distorto

$$\left[ \text{Risposta } MSE[T_1] = \frac{(b+1)^2}{24} \right]$$

$$MSE[T_2] = \frac{(b+1)^2}{6}$$

$T_1$  preferibile

Esercizio 14 (Tema d'esame del 05/04/2005).

Sia  $X$  una variabile casuale normale con media  $\mu$  e varianza  $\sigma^2$ . Siano  $X_1, X_2, X_3$  le variabili casuali indipendenti descritte dalle tre determinazioni  $x_1, x_2, x_3$  di un campione casuale estratto da essa. Per stimare il parametro  $\mu$  si considerano i due seguenti stimatori:

$$\bar{X}_3 = \frac{X_1 + X_2 + X_3}{3}, \quad T = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{3}{5}X_3.$$

1. Dire se  $\bar{X}_3$  e  $T$  sono stimatori non distorti di  $\mu$  e motivare la risposta.
2. Calcolare  $MSE[\bar{X}_3]$  e  $MSE[T]$  e stabilire quale tra i due stimatori  $\bar{X}_3$  e  $T$  di  $\mu$  sia preferibile, motivando la risposta.

$\bar{X}_3$  non distorto

$T$  non distorto

[Risposta  $MSE[\bar{X}_3] = \frac{\sigma^2}{3}$  ]

$MSE[T] = \frac{11\sigma^2}{25}$

$\bar{X}_3$  preferibile