UNIVERSITÀ DEGLI STUDI ROMA TRE
Corso di Laurea in Matematica
ST410 - Statistica 1 - A.A. 2014/2015

I Esonero - 6 Novembre 2014

1	2	3	4	5	6	7	8	Tot.

Avvertenza: Svolgere ogni esercizio nello spazio assegnato, senza consegnare altri fogli e accompagnando le risposte con spiegazioni complete, chiare ed essenziali. Scrivere il proprio nome su ogni foglio nello spazio predisposto. Non è consentito l'utilizzo di libri o appunti; non è consentito l'uso di calcolatrici.
Tempo: 2h30.

COGNOME: .... **NOME:** ....

MATRICOLA: ....

Nome e Cognome:

Esercizio 1. Sia Z variabile aleatoria con densità di probabilità triangolare

$$f_Z(z) = \begin{cases} z & \text{se } 0 \leq z \leq 1 \\ 2-z & \text{se } 1 \leq z \leq 2 \\ 0 & \text{altrimenti} \end{cases}$$

1. (2pt) Calcolare $m_Z(t)$, la funzione generatrice dei momenti di Z. Quanto vale $m_Z(0)$?
2. (2pt) Siano X, Y variabili aleatorie uniformi su $[0,1]$ e indipendenti. Mostrare che $W = X+Y$ si distribuisce come Z.

Soluzione:

1. $m_Z(t) = E(e^{tZ})$ (se esiste tale media in un intorno di 0). Quindi

   $$E(e^{tZ}) = \int_{-\infty}^{+\infty} e^{tz} f_Z(z)dz = \int_0^1 e^{tz} z dz + \int_1^2 e^{tz}(2-z)dz.$$

   Integrando per parti, il primo addendo è pari a $\frac{te^t+1-e^t}{t^2}$, il secondo a $\frac{-te^t+e^{2t}-e^t}{t^2}$. Quindi, per definizione, $m_Z(0) = 1$, mentre $m_Z(t) = \frac{(e^t-1)^2}{t^2}$ per ogni $t \neq 0$.
2. Calcoliamo la funzione generatrice dei momenti di X (e quindi di Y): $m_X(t) = E(e^{tX}) = \int_0^1 e^{tx}dx = \frac{e^t-1}{t}$ per ogni $t \neq 0$ (e $m_X(0) = 1$). Dato che X e Y sono indipendenti, allora $m_W(t) = m_X(t) \cdot m_Y(t) = \frac{(e^t-1)^2}{t^2}$. Dato che $m_W(t) = m_Z(t)$ per ogni t in un intorno di 0 (in questo caso per ogni $t \in \mathbb{R}$), allora W e Z sono indenticamente distribuite.

Nome e Cognome:

Esercizio 2. Sia (X, Y) vettore casuale con densità congiunta

$$f_{X,Y}(x, y) = \frac{e^{-y}}{y} \mathrm{I}_{(0,y)}(x) \mathrm{I}_{(0,+\infty)}(y),$$

dove I_A è la funzione indicatrice dell'intervallo $A \subseteq \mathbb{R}$.

1. (1pt) Calcolare la densità marginale $f_Y(y)$.
2. (1pt) Calcolare $E(X^2|Y = y)$ per ogni $y > 0$.
3. (1pt) Dato il punto precedente, determinare $E(X^2|Y)$.
4. (2pt) Quanto vale $E(E(X^2|Y))$?

Soluzione:

1. Per $y \leq 0$, $f_Y(y) = 0$; per ogni $y \in (0, +\infty)$, $f_Y(y) = \int_0^y \frac{e^{-y}}{y} dx = e^{-y}$.
2. La densità condizionata di X dato $Y = y$ (con $y > 0$) è

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{y} \mathrm{I}_{(0,y)}(x).$$

   Perciò $h(y) := E(X^2|Y = y) = \int_0^y x^2 \frac{1}{y} dx = \frac{y^2}{3}$.
3. Per definizione $E(X^2|Y) = h(Y) = \frac{Y^2}{3}$.
4. $E\left(\frac{Y^2}{3}\right) = \int_0^{+\infty} \frac{y^2}{3} f_Y(y)dy = \frac{1}{3} \int_0^{+\infty} y^2 e^{-y} dy = \frac{1}{3}\Gamma(3)$, dove $\Gamma(t)$ è la funzione gamma. Siccome $\Gamma(3) = 2! = 2$, allora $E(E(X^2|Y)) = \frac{2}{3}$.
   Si può ottenere la stessa soluzione anche ricordandosi che, per quanto visto a lezione, $E(E(X^2|Y)) = E(X^2)$. In questo caso è necessario calcolarsi (per ogni $x > 0$) la densità marginale $f_X(x) = \int_{-\infty}^{+\infty} \frac{e^{-y}}{y} \mathrm{I}_{(0,y)}(x) \mathrm{I}_{(0,+\infty)}(y) dy$. Dato che, fissato $x > 0$, $\mathrm{I}_{(0,y)}(x) = \mathrm{I}_{(x,+\infty)}(y)$, allora $f_X(x) = \int_x^{+\infty} \frac{e^{-y}}{y} dy$, quindi $E(X^2) = \int_0^{+\infty} x^2 \left(\int_x^{+\infty} \frac{e^{-y}}{y} dy\right) dx = \int_0^{+\infty} \int_x^{+\infty} x^2 \frac{e^{-y}}{y} dydx = \int_0^{+\infty} \int_0^y x^2 \frac{e^{-y}}{y} dxdy = \int_0^{+\infty} \frac{e^{-y}}{y} \frac{y^3}{3} dy = \frac{1}{3} \int_0^{+\infty} y^2 e^{-y} dy = \frac{2}{3}$.

Nome e Cognome:

Esercizio 3. Sia $f(x)$ densità con media μ e varianza finita σ^2 e sia $X_1, \ldots, X_n$ un campione casuale di ampiezza n estratto da $f(x)$. Si indichi con $\overline{X}_n$ la media campionaria.

1. (2pt) Si enunci la legge debole dei grandi numeri.
2. (2pt) Si enunci il teorema del limite centrale.
3. (2pt) Quanto deve essere grande n affinché $\mathrm{P}\left(|\overline{X}_n - \mu| < \frac{\sigma}{5}\right) \geq 95\%$? Si risponda utilizzando sia la legge debole dei grandi numeri che il teorema del limite centrale.

   Può essere utile conoscere il valore di uno dei seguenti quantili notevoli di una normale standardizzata (approssimati al primo decimale): quantile di ordine 0.95: 1.6; quantile di ordine 0.975: 2.0; quantile di ordine 0.99: 2.3.

Soluzione:

1. Cfr. Mood par. 6.3.2
2. Cfr. Mood par. 6.3.3
3. Utilizziamo la notazione di Mood.

   Per la legge debole dei grandi numeri se $n \geq \frac{\sigma^2}{\epsilon^2 \delta}$ allora $\mathrm{P}(|\overline{X}_n - \mu| < \epsilon) \geq 1 - \delta$. Nel nostro caso $\delta = 0.05$, $\epsilon = \frac{\sigma}{5}$, quindi $n \geq 500$.

   Per il teorema del limite centrale, se n è sufficientemente grande $\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ si distribuisce come una normale standardizzata. Perciò

   $$\mathrm{P}\left(|\overline{X}_n - \mu| < \frac{\sigma}{5}\right) = \mathrm{P}\left(\left|\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}\right| < \frac{\sqrt{n}}{5}\right) \cong \mathrm{P}\left(\frac{-\sqrt{n}}{5} \leq Z \leq \frac{\sqrt{n}}{5}\right),$$

   dove Z è una normale standardizzata. Dato che si vuole che tale probabilità sia maggiore di 95%, allora si deve scegliere n tale che $\frac{\sqrt{n}}{5} \geq q_{0.975}$, dove $q_{0.975}$ è il quantile di ordine 0.975 di Z. Dato che $q_{0.975} \cong 2.0$ otteniamo $n \geq 100$.

Nome e Cognome:

Esercizio 4. Sia $X_1, \ldots, X_n$ un campione casuale di ampiezza $n \geq 2$ estratto da una popolazione con densità $f(x)$ con media μ e varianza finita σ^2. Si indichi con $\overline{X}$ la media campionaria e con S^2 la varianza campionaria.

1. (2pt) Dimostrare che $E(S^2) = \sigma^2$.
2. (2pt) Nel caso in cui la densità sia quella di una normale calcolare $\mathrm{Var}(\overline{X} + S^2)$.

Soluzione:

1. Cfr. Mood par. 6.2.4 o lezione del 7 ottobre.
2. Per quanto visto a lezione, nel caso di una densità normale $\overline{X}$ e S^2 sono indipendenti. Quindi $\mathrm{Var}(\overline{X} + S^2) = \mathrm{Var}(\overline{X}) + \mathrm{Var}(S^2)$. Ora: $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$, mentre $\frac{n-1}{\sigma^2} S^2$ si distribuisce come una chi-quadro con $n-1$ gradi di libertà e quindi la sua varianza è $2(n-1)$, perciò $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}$. Quindi $\mathrm{Var}(\overline{X} + S^2) = \frac{\sigma^2}{n} + \frac{2\sigma^4}{n-1}$.

Nome e Cognome:

Esercizio 5. Siano X_1, X_2, X_3, X_4 variabili aleatorie normali indipendenti e identicamente distribuite, di media 0 e varianza σ^2.

1. (2pt) Determinare la distribuzione e la media di $\frac{\sqrt{2}X_1}{\sqrt{X_2^2+X_3^2}}$.

2. (2pt) Determinare la densità di probabilità di $X_1^2 + X_2^2 + 1$.

Soluzione:

Per $i = 1, \dots, 4$, sia $X_i' = \frac{X_i}{\sigma}$. Per ogni $i = 1, \dots, 4$ X_i' è una normale standardizzata.

1. $\frac{\sqrt{2}X_1}{\sqrt{X_2^2+X_3^2}} = \frac{X_1'}{\sqrt{(X_2'^2+X_3'^2)/2}}$, con X_1' normale standardizzata, indipendente da $\left(X_2'^2 + X_3'^2\right)/2$. Dato che X_2' e X_3' sono indipendenti, allora $X_2'^2 + X_3'^2 \sim \chi_2^2$, perciò, per quanto visto a lezione, $\frac{\sqrt{2}X_1}{\sqrt{X_2^2+X_3^2}}$ si distribuisce come una t di Student con due gradi di libertà e la sua media è 0.

2. $W = X_1^2 + X_2^2 + 1 = \sigma^2(X_1'^2 + X_2'^2) + 1$, con $Y = X_1'^2 + X_2'^2 \sim \chi_2^2$ (X_1' e X_2' sono indipendenti). Quindi $f_Y(y) = \frac{1}{2}e^{-y/2} \cdot \mathrm{I}_{(0,+\infty)}(y)$ (esponenziale di parametro $\frac{1}{2}$).

   Sia $g : \mathbb{R} \to \mathbb{R}, g(y) := \sigma^2 y + 1$. g è lineare e $W = g(Y)$. L'inversa $g^{-1} : \mathbb{R} \to \mathbb{R}$ è definita come $g^{-1}(w) = \frac{w-1}{\sigma^2}$. Applicando il teorema di cambio di variabili (le ipotesi sono soddisfatte perché g è lineare) abbiamo che:

$$f_W(w) = f_Y(g^{-1}(w))|(g^{-1})'(w)| = \frac{1}{2\sigma^2}e^{-(w-1)/(2\sigma^2)} \cdot \mathrm{I}_{(1,+\infty)}(w).$$

   Oppure: $F_W(w) = P(W \leq w) = P(\sigma^2 Y + 1 \leq w) = P(Y \leq (w-1)/\sigma^2) = F_Y((w-1)/\sigma^2)$. Quindi $f_W(w) = F_W(w)' = (F_Y((w-1)/\sigma^2))' = f_Y((w-1)/\sigma^2)) \cdot \frac{1}{\sigma^2}$. Da cui

$$f_W(w) = \frac{1}{2\sigma^2}e^{-(w-1)/(2\sigma^2)} \cdot \mathrm{I}_{(1,+\infty)}(w).$$

Nome e Cognome:

Esercizio 6. Si consideri un campione casuale $\underline{X} = (X_1, \ldots, X_n)$ da una popolazione di densità $f(x;\theta)$, con $x \in \mathbb{R}$ e $\theta \in \Theta \subseteq \mathbb{R}$. Siano $S(\underline{X}), T(\underline{X})$ due statistiche, con $S, T : \mathbb{R}^n \to \mathbb{R}$.

1. (3pt) Si definisca cosa vuol dire che S è una statistica sufficiente per θ e cosa vuol dire che S è una statistica sufficiente e minimale per θ.

2. (2pt) Sia $f(x;\theta) = (\theta+1)\theta x^{\theta-1}(1-x)\mathrm{I}_{(0,1)}(x)$, con $\Theta = (0, +\infty)$ e $\mathrm{I}_{(0,1)}$ funzione indicatrice dell'intervallo $(0,1) \subsetneq \mathbb{R}$. Dimostrare che $W = \Pi_{i=1}^n X_i$ è una statistica sufficiente e minimale per θ.

Soluzione:

1. Cfr. Mood par. 7.4.1,7.4.3 o lezioni del 21 e 23 ottobre.

2. Utilizziamo il principio di sufficienza minimale di Lehmann-Scheffé. Scriviamo la densità congiunta del campione: per ogni $\underline{x}$ nello spazio campionario $\chi = (0,1)^n$
$$f_{\underline{X}}(\underline{x};\theta) = (\theta+1)^n\theta^n\Pi_{i=1}^n(1-x_i)\left(\Pi_{i=1}^n x_i\right)^{\theta-1}.$$
Per ogni $\underline{x}, \underline{y} \in \chi$,
$$\frac{f_{\underline{X}}(\underline{x};\theta)}{f_{\underline{X}}(\underline{y};\theta)} = \frac{\Pi_{i=1}^n(1-x_i)}{\Pi_{i=1}^n(1-y_i)}\left(\frac{\Pi_{i=1}^n x_i}{\Pi_{i=1}^n y_i}\right)^{\theta-1}.$$
Per ogni $\underline{x}, \underline{y} \in \chi$ tale rapporto non dipende da θ se e solo se $\left(\frac{\Pi_{i=1}^n x_i}{\Pi_{i=1}^n y_i}\right) = 1$, i.e., se e solo se $W(\underline{x}) = W(\underline{y})$. Quindi W è una statistica sufficiente e minimale per θ.

Nome e Cognome:

Esercizio 7. Sia $X_1, \ldots, X_n$ un campione casuale estratto da una popolazione di densità $f(x;\theta) = \theta x^{\theta-1} \mathrm{I}_{(0,1)}(x)$ con $\theta > 0$ e $\mathrm{I}_{(0,1)}$ funzione indicatrice dell'intervallo $(0,1) \subsetneq \mathbb{R}$.

1. (1pt) Determinare uno stimatore di θ con il metodo dei momenti.

2. (2pt) Determinare uno stimatore di θ con il metodo della massima verosimiglianza.

Soluzione:

1. Uguagliamo la media campionaria alla media della popolazione. Sia X v.a. con densità $f(x;\theta)$. Allora $E(X) = \int_0^1 \theta x^\theta dx = \frac{\theta}{\theta+1}$. La media campionaria coincide con il momento assoluto di ordine 1, quindi è pari a $\overline{X} := \frac{1}{n}\sum_{i=1}^n X_i$. Perciò $\frac{\theta}{\theta+1} = \overline{X}$, da cui: uno stimatore $\hat{\theta}$ di θ ottenuto con il metodo dei momenti è $\hat{\theta}(X_1, \ldots, X_n) = \frac{\overline{X}}{1-\overline{X}}$.

2. Scriviamo la funzione di verosimiglianza del campione: per ogni $\underline{x} \in (0,1)^n$,

$$L(\theta, \underline{x}) = \theta^n \left(\Pi_{i=1}^n x_i\right)^{\theta-1}.$$

Prendendo il logaritmo:

$$\log(L(\theta, \underline{x})) = n\log(\theta) + (\theta - 1)\sum_{i=1}^n \log(x_i).$$

Derivando la funzione di log-verosimiglianza si ottiene:

$$\frac{d}{d\theta}\log(L(\theta, \underline{x})) = \frac{n}{\theta} + \sum_{i=1}^n \log(x_i),$$

quindi la funzione di log-verosimiglianza ha un unico punto critico, che è un massimo locale, per $\theta = -\frac{n}{\sum_{i=1}^n log(x_i)}$. Non essendoci altri punti critici il massimo locale è necessariamente anche un massimo globale, quindi uno stimatore con il metodo della massima verosimiglianza è

$$\hat{\theta}(X_1, \ldots, X_n) = -\frac{n}{\sum_{i=1}^n \log(X_i)}.$$

Nome e Cognome:

Esercizio 8. (3pt) Sia $\underline{X} = (X_1, \ldots, X_n)$ campione casuale da una popolazione con densità $f(x;\theta)$ con $x \in \mathbb{R}$ e $\theta \in \Theta \subseteq \mathbb{R}$. Sia $L(\theta;\underline{x})$ (con $\underline{x} = (x_1, \ldots, x_n)$) la funzione di verosimiglianza del campione. Sia $S(\underline{X})$ statistica sufficiente, con S funzione da $\mathbb{R}^n$ in $\mathbb{R}$. Si supponga che per ogni $\underline{x} \in \mathbb{R}^n$ esista uno e un solo $\hat{\theta}(\underline{x})$ che massimizza $L(\theta;\underline{x})$. Dimostrare che lo stimatore di massima verosimiglianza $\hat{\theta}(\underline{X})$ è funzione di $S(\underline{X})$ (i.e.: se $S(\underline{x}) = S(\underline{y})$ allora $\hat{\theta}(\underline{x}) = \hat{\theta}(\underline{y})$).

Soluzione:

Utilizziamo il teorema di fattorizzazione di Neyman–Fisher: dato che S è sufficiente per θ allora esistono funzioni positive $h : \mathbb{R}^n \to \mathbb{R}$ e $g : \mathbb{R} \to \mathbb{R}$ (con g eventualmente dipendente da θ) tali che la distribuzione congiunta di $\underline{X}$, $f_{\underline{X}}(\underline{x};\theta)$, si fattorizza come

$$f_{\underline{X}}(\underline{x};\theta) = h(\underline{x})g(S(\underline{x});\theta).$$

Quindi, fissato $\underline{x} \in \mathbb{R}^n$,

$$L(\theta;\underline{x}) = h(\underline{x})g(S(\underline{x});\theta).$$

Perciò il il valore θ che massimizza $L(\theta;\underline{x})$ dipende esclusivamente da $g(S(\underline{x});\theta)$. Inoltre, per ipotesi, tale valore è unico. Quindi se $S(\underline{x}) = S(\underline{y})$ allora $\hat{\theta}(\underline{x}) = \hat{\theta}(\underline{y})$.