UNIVERSITÀ DEGLI STUDI ROMA TRE
Corso di Laurea in Matematica
ST410 - Statistica 1 - A.A. 2013/2014

I Esonero - 29 Ottobre 2013

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; è consentito l'uso di una calcolatrice non programmabile.

Io sottoscritto/a

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

MATRICOLA: ....

dichiaro, sotto la mia responsabilità, di aver svolto l'odierna prova di esame in modo autonomo e senza l'ausilio di strumenti o materiale non permesso.

Roma, 29 Ottobre 2013

FIRMA

Esercizio 1. Sia X una variabile casuale continua con densità di probabilità $f_X(\cdot)$.

1. (1pt) Si definisca la *funzione generatrice dei momenti di* X, $m_X(t)$.
2. (1pt) Supponendo che per X esista la funzione generatrice dei momenti, dimostrare che per ogni a, b costanti in $\mathbb{R}$, $m_{aX+b}(t) = e^{bt} m_X(at)$.
3. (2pt) Supponendo che f_X sia una funzione pari (i.e., $f_X(x) = f_X(-x)$ per ogni $x \in \mathbb{R}$) dimostrare che X e $-X$ hanno la stessa distribuzione e che $m_X(t)$ (se esiste) è una funzione pari in t.

Soluzione:

1. Cfr. , ad esempio, Mood - Graybill - Boes - "Introduzione alla statistica" - McGraw-Hill - pag. 88
2. $m_{aX+b}(t) := E(e^{(aX+b)t}) = e^{bt}E(e^{atX})$. Per $|t|$ suff. piccolo $E(e^{atX})$ esiste per ipotesi, e per definizione vale $m_X(at)$.
3. Si può vedere in più modi: ad esempio, applicando il teorema di trasformazione, dato che la funzione $g(x) = -x$ è invertibile su tutto $\mathbb{R}$. Notiamo che $|g'(x)| = 1$, quindi $f_{-X}(x) = f_X(-x) = f(x)$, quest'ultima uguaglianza per ipotesi. Inoltre $m_X(-t) = E(e^{-tX}) = E(e^{t(-X)})$. Dato che X e $-X$ hanno la stessa distribuzione, $E(e^{t(-X)}) = E(e^{tX})$.

Nome e Cognome:

Esercizio 2. Siano X Y variabili casuali continue con funzione di densità di probabilità congiunta

$$f_{X,Y}(x,y) = \begin{cases} 9x^2y^2 & \text{se } 0 \leq x \leq 1 \text{ e } 0 \leq y \leq 1 \\ 0 & \text{altrimenti} \end{cases}$$

1. (2pt) Calcolare $P\left(XY \leq \frac{1}{3}, X \geq \frac{1}{3}\right)$.
2. (2pt) Stabilire se X e Y sono stocasticamente indipendenti.
3. (1pt) Calcolare $E(XY)$.

Soluzione:

1. Sia $A := \{xy \leq 1/3, x \geq 1/3, 0 \leq x \leq 1, 0 \leq y \leq 1\} \subseteq \mathbb{R}^2$ Allora $P\left(XY \leq \frac{1}{3}, X \geq \frac{1}{3}\right) = \iint_A 9x^2y^2 dxdy = \int_{1/3}^1 dx \left(\int_0^{1/(3x)} 9x^2y^2 dy\right) = \int_{1/3}^1 \frac{1}{9x} dx = \frac{1}{9}\log(3)$.
2. Per controllare se X e Y sono stocasticamente indipendenti basta calcolare le funzioni di densità marginali $f_X(\cdot)$ e $f_Y(\cdot)$ e verificare se $f_{X,Y} = f_X \cdot f_Y$.
   $f_X(x) = \int_0^1 9x^2y^2 dy = 3x^2$. Analogamente $f_Y(y) = 3y^2$. Ma allora $f_{X,Y} = f_X \cdot f_Y$ e quindi X e Y sono indipendenti.
3. $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 3x^3 dx = \frac{3}{4}$. Analogamente $E(Y) = \frac{3}{4}$. Dato che X e Y sono indipendenti $E(XY) = E(X)E(Y) = \frac{9}{16}$.

Esercizio 3. Sia X una variabile casuale continua con funzione di densità di probabilità $f_X(\cdot)$ e sia $X_1, \ldots, X_n$ un campione casuale di ampiezza n estratto da una popolazione con la stessa legge di X.

1. (1pt) Definire cosa si intende per mediana della distribuzione di X.
2. (1pt) Definire cosa si intende per campione casuale di ampiezza n.
3. (1pt) Qual è la probabilità che la più grande tra $X_1, \ldots, X_n$ sia uguale o maggiore della mediana della popolazione?

Soluzione:

1. Cfr., ad esempio, Mood - Graybill - Boes - "Introduzione alla statistica" - McGraw-Hill - pag. 84
2. Cfr., ad esempio, Mood - Graybill - Boes - "Introduzione alla statistica" - McGraw-Hill - pag. 230
3. Sia ξ tale che $F_X(\xi) = 0.5$. $P(\max\{X_1 \ldots X_n\} \geq \xi) = 1 - P(\max\{X_1 \ldots X_n\} < \xi) = 1 - P(X_1 < \xi, \ldots, X_n < \xi)$. Dato che $X_1, \ldots, X_n$ è un campione casuale allora $P(X_1 < \xi, \ldots, X_n < \xi) = P(X_1 < \xi) \cdot \cdots \cdot P(X_n < \xi)$. Ora: $P(X_i < \xi) = 0.5$ per definizione, quindi la probabilità cercata è $1 - \frac{1}{2^n}$.

Nome e Cognome:

Esercizio 4. Siano X e $X_1, X_2, \dots$ variabili casuali i.i.d. con distribuzione esponenziale di parametro $\lambda = 1$. Sia $g : \mathbb{R} \to \mathbb{R}$ tale che $g(x) = 1$ se $1 \leq x \leq 2$, 0 altrimenti.

1. (2pt) Enunciare la legge debole dei grandi numeri.
2. (1pt) Calcolare $a := P(1 \leq X \leq 2)$.
3. (2pt) Siano $Y_i = g(X_i)$ per ogni $i \geq 1$.
   Per ogni $n \geq 1$ determinare la distribuzione di $W_n := \sum_{i=1}^n Y_i$.
   Sia $\overline{Y}_n = \frac{1}{n} W_n$: calcolarne la media.
   Per ogni $\epsilon > 0$ calcolare $\lim_{n \to +\infty} P(|\overline{Y}_n - a| < \epsilon)$

Soluzione:

1. Cfr. ad esempio, Mood - Graybill - Boes - "Introduzione alla statistica" - McGraw-Hill - pag. 240.
2. $f_X(x) = e^{-x}\mathbb{I}_{[0,\infty)}$, quindi $a = \int_1^2 f_X(x)dx = e^{-1} - e^{-2}$.
3. Y_i assume valore 1 se $1 \leq X_i \leq 2$, quindi con probabilità a; altrimenti è pari a 0. Quindi Y_i è una bernoulliana di parametro $p = a$. Dato che $X_1, \dots, X_n$ sono indipendenti, allora anche $Y_1, \dots, Y_n$ sono indipendenti, perciò W è una binomiale di parametri a e n.
   $E(\overline{Y}_n) = E(X) = a$.
   Per la legge debole dei grandi numeri il limite è 1.

Esercizio 5. Un campione casuale $X_1, \dots, X_{12}$ è estratto da una distribuzione uniforme sull'intervallo $[0, 1]$.

1. (2pt) Enunciare il teorema limite centrale.
2. (3pt) Utilizzando il teorema limite centrale trovare un'approssimazione per $P\left(\left|\overline{X}_{12} - \frac{1}{2}\right| \leq 0.1\right)$, dove $\overline{X}_{12}$ è la media campionaria.

Soluzione:

1. Cfr., ad esempio, Mood - Graybill - Boes - "Introduzione alla statistica" - McGraw-Hill - pag. 241
2. $P\left(\left|\overline{X}_{12} - \frac{1}{2}\right| \leq 0.1\right) = P\left(\left|\overline{X}_{12} - \frac{1}{2}\right| / \frac{1}{12} \leq 1.2\right)$.
   La varianza di una distribuzione uniforme su $[0, 1]$ è $\frac{1}{12}$, quindi per il teorema limite centrale $P\left(\left|\overline{X}_{12} - \frac{1}{2}\right| / \frac{1}{12} \leq 1.2\right) \approx P(|Z| < 1.2)$, dove $Z \sim N(0, 1)$.

   $$P(|Z| < 1.2) = P(-1.2 \leq Z \leq 1.2) = 2P(Z \leq 1.2) - 1 \approx 0.7698.$$

   L'ultima equivalenza si desume dalla tavola della funzione di ripartizione di una normale standardizzata.

Nome e Cognome:

Esercizio 6. (3pt) Siano X e Y due variabili casuali i.i.d $\sim N(0,1)$. Sia $Z := \min\{X, Y\}$. Dimostrare che Z^2 ha una distribuzione chi-quadrato con un grado di libertà.

Soluzione:

Sia F la funzione di ripartizione di una normale standardizzata N.
In generale $P(Z \leq a) = 1 - P(Z > a) = 1 - P(X > a, Y > a)$. Data l'indipendenza di X e Y, $1 - P(X > a, Y > a) = 1 - P(X > a)P(Y > a) = 1 - (1 - F(a))(1 - F(a)) = 2F(a) - F(a)^2$.
Ora: $P(Z^2 \leq a^2) = P(-a \leq Z \leq a) = P(Z \leq a) - P(Z < -a) = 2F(a) - F(a)^2 - 2F(-a) + F(-a)^2$.
Per le note proprietà di F otteniamo quindi che $P(Z^2 \leq a^2) = 2F(a) - 1 = P(-a \leq N \leq a) = P(N^2 \leq a^2)$. N^2, essendo il quadrato di una normale standardizzata, si comporta come una chi-quadrato con un grado di libertà.

Esercizio 7. Sia X_i una variabile casuale con distribuzione $N(i, i^2)$, per $i = 1, 2, 3$. Assumete che X_1, X_2, X_3 siano indipendenti.

1. (1pt) Qual è la distribuzione di $X_1 + X_2 + X_3$?
2. (1pt) Qual è la distribuzione di $\frac{\sqrt{5}X_1+X_2-2-\sqrt{5}}{\sqrt{(X_3-3)^2}}$?
3. (1pt) Utilzzando solo le tre variabili X_1, X_2, X_3 fare un esempio di statistica avente una distribuzione chi-quadrato con 3 gradi di libertà.
4. (1pt) Utilizzando solo le tre variabili X_1, X_2, X_3 fare un esempio di statistica avente una distribuzione t di Student con due gradi di libertà.

Soluzione:

1. Una combinazione lineare di variabili normali indipendenti è ancora una variabile normale. Quindi per conoscere la distribuzione di $X_1 + X_2 + X_3$ basta calcolarne media e varianza. $E(X_1+X_2+X_3) = E(X_1)+E(X_2)+E(X_3) = 1+2+3 = 6$. Data l'indipendenza: $Var(X_1 + X_2 + X_3) = Var(X_1) + Var(X_2) + Var(X_3) = 14$. Quindi $X_1 + X_2 + X_3 \sim N(6, 14)$.
2. La variabile al numeratore è una gaussiana di media 0 e varianza 9. Anche (X_3-3) è una gaussiana di media 0 e varianza 9, indipendente dalla prima. Perciò:$\left(\frac{X_3-3}{3}\right)^2$ è una chi-quadrato con un grado di libertà, indipendente dal numeratore. Inoltre $\frac{\sqrt{5}X_1+X_2-2-\sqrt{5}}{3}$ è una normale standardizzata. Ma allora il rapporto cercato è equivalente al rapporto tra una normale standardizzata e la radice quadrata di una variabile chi-quadrato, indipendente dalla prima e divisa per i suoi gradi di libertà (nella fattispecie 1). Perciò il rapporto ha una distribuzione t di Student con un grado di libertà.
3. La somma dei quadrati di tre normali standardizzate ed indipendenti ha una distribuzione chi-quadrato con 3 gradi di libertà. Quindi un possibile esempio è $(X_1 - 1)^2 + \left(\frac{X_2-2}{2}\right)^2 + \left(\frac{X_3-3}{3}\right)^2$.
4. Il rapporto tra una normale standardizzata e la radice della metà di una chi-quadrato, indipendente dalla prima, ha distribuzione t di Student con 2 gradi di libertà. Un esempio quindi è
$$\frac{X_1 - 1}{\sqrt{\left(\frac{X_2-2}{2}\right)^2 + \left(\frac{X_3-3}{3}\right)^2}}.$$

Nome e Cognome:

Esercizio 8. Sia $X_1, \ldots, X_n$ un campione casuale di ampiezza n estratto da una popolazione con distribuzione uniforme nell'intervallo $[a-b, a+b]$ $(b > 0)$.

1. (1pt) Scrivere la funzione di densità di probabilità della distribuzione uniforme nell'intervallo $[a-b, a+b]$.

2. (3pt) Determinare gli stimatori di a e b con il metodo dei momenti.

Soluzione:

Sia X v.a. con distribuzione uniforme nell'intervallo $[a-b, a+b]$.

1. $f_X(x) = \frac{1}{(a+b)-(a-b)} \mathbb{I}_{[a-b,a+b]} = \frac{1}{2b} \mathbb{I}_{[a-b,a+b]}$.

2. $E(X) = a$, quindi lo stimatore di a è la media campionaria $\overline{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$.
   $Var(X) = \frac{b^2}{3}$, quindi $\frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2$ è uno stimatore di $b^2/3$. Perciò

$$\sqrt{3\left(\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2\right)} = \sqrt{3\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2\right)}$$

   è uno stimatore di b.

(eventuale continuazione esercizio n.:)

Tabella 1: Funzione di ripartizione di una normale standardizzata

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000