Corso di Laurea in Ingegneria Informatica
Anno Accademico 2021/2022
Calcolo delle Probabilità e Statistica Matematica

Nome ....

N. Matricola .... Ancona, 9 febbraio 2022

1. Siano $X_1, X_2, ..., X_n$ variabili aleatorie indipendenti. Siano σ_j^2 e μ_j le loro varianze e le loro medie e siano ψ_j i loro momenti centrati del terz'ordine. Sia $Y = X_1 + X_2 + ... + X_n$ la loro somma e indichiamo con μ e σ^2 la media e la varianza di Y. Sia inoltre ψ il momento centrato del terz'ordine di Y.

   - Dimostrare che $\psi = \psi_1 + \psi_2 + ... + \psi_n$;
   - dimostrare che la proprietà additiva non vale per i momenti centrati del quart'ordine e superiori.

   (è sufficiente fare la dimostrazione per $n = 2$).

Soluzione.

$$\psi_j = E[(X_j - \mu_j)^3]$$

$$\psi = E[(Y - \mu_Y)^3] \qquad \mu_Y = \sum_j \mu_j$$

$$Y - \mu_Y = \sum_j (X_j - \mu_j) = (X_1 - \mu_1) + (X_2 - \mu_2) + \cdots$$

$$\psi = E[(Y - \mu_Y)^3] = E\Big[(X_1 - \mu_1)^3 + \cdots + (X_n - \mu_n)^3 +$$

$$\sum_{j=1}^{n} (X_j - \mu_j) \prod_{k,l \neq j} (X_k - \mu_k)(X_l - \mu_l)\Big] =$$

$$= \psi_1 + \psi_2 + \dots + \psi_n +$$

$$+ \underbrace{\sum_{j=1}^{n} E[X_j - \mu_j]}_{=0} \prod_{k,l \neq j} E[(X_k - \mu_k)(X_l - \mu_l)]$$

$$= \sum_j \psi_j + \text{zero}$$

- $E[(Y - \mu_Y)^n] \qquad n \geq 4$

$\neq \sum_j [(X_j - \mu_j)^n]$ perché non tutti i termini contengono $E[X_j - \mu_j]$ linearmente

2. Si estrae un numero X a caso nell'intervallo $(0,1)$ ed un secondo numero Y nell'intervallo $(0,X]$. Si sa che la distribuzione congiunta di X e Y è data da

$$f_{XY} = \frac{1}{x} \quad \text{se} \quad 0 < x < 1, 0 < y < x$$
$$= 0 \quad \text{altrimenti.}$$

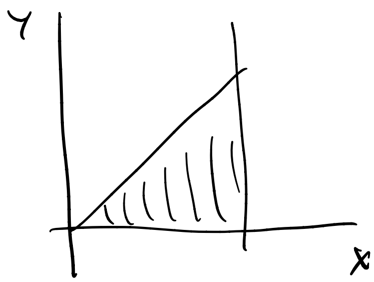
(a) Dimostrare che

$$E[X^r Y^s] = \frac{1}{(s+1)\,(r+s+1)}, \qquad r,s \in \mathbb{N};$$

(b) verificare la formula in (a) per $r = 2$ ed $s = 0$, calcolando direttamente $E[X^2]$;

(c) usando la formula in (a), calcolare il coefficiente di correlazione di X e Y.

Soluzione.



Verifica: $\int_0^1 dx \int_0^x dy \, \frac{1}{x} = \int_0^1 dx \, \frac{1}{x} \, x = 1$ $\underline{\underline{\text{ok}}}$

Continua

(a) $E[X^r Y^s] = \int_0^1 dx \int_0^x dy \, \frac{1}{x} x^r y^s =$

$$= \int_0^1 dx \, x^{r-1} \int_0^x dy \, y^s =$$

$$= \int_0^1 dx \, x^{r-1} \frac{x^{s+1}}{s+1} = \frac{1}{s+1} \int_0^1 dx \, x^{r+s} =$$

$$= \frac{1}{s+1} \left. \frac{x^{r+s+1}}{r+s+1} \right|_0^1 = \frac{1}{(s+1)(r+s+1)}$$

(b) $E[X^2] = \int_0^1 dx \int_0^x dy \, \frac{1}{x} x^2 =$

$$= \int_0^1 dx \, x^2 = \frac{1}{3}$$

$\frac{1}{(s+1)(r+s+1)}$ per $r=2$ e $s=0$: $\frac{1}{3}$

(c) $$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

$$\mathrm{Cov}(X,Y) = \underset{r=s=1}{E[XY]} - \underset{\substack{r=1\\ s=0}}{E[X]}\,\underset{\substack{r=0\\ s=1}}{E[Y]} =$$

$$= \frac{1}{2\cdot 3} - \frac{1}{2}\cdot\frac{1}{4} = \frac{1}{6} - \frac{1}{8} = \frac{1}{24}$$

$$\mathrm{Var}(X) = \underset{\substack{r=2\\ s=0}}{E[X^2]} - \underset{\substack{r=1\\ s=0}}{E[X]}^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

$$\mathrm{Var}(Y) = \underset{\substack{r=0\\ s=2}}{E[Y^2]} - \underset{\substack{r=0\\ s=1}}{E[Y]}^2 = \frac{1}{9} - \frac{1}{16} = \frac{7}{144}$$

$$\rho = \frac{1/24}{\sqrt{(1/12)(7/144)}} = \frac{1}{24}\sqrt{\frac{12\cdot 144}{7}} = \frac{1}{2}\sqrt{\frac{12}{7}} \approx$$

$$\approx 0.65$$

3. Si vuole determinare l'intervallo di confidenza per la densità dei portatori di carica in un campione di silicio. Si fanno due misurazioni che danno luogo ai campioni di rango $n = 10$: (in unità di 10^{10} cm^{-3}):

$$I : 1.34, 1.47, 1.48, 1.41, 1.38, 1.35, 1.39, 1.47, 1.42, 1.48$$
$$II : 1.44, 1.48, 1.47, 1.50, 1.46, 1.54, 1.46, 1.48, 1.48, 1.43$$

Determinare, per entrambi i campioni, gli intervalli di fiducia al 90%, 95% e 99%. In quali casi gli intervalli provenienti dai due campioni si sovrappongono?

Soluzione. Abbiamo

$$\begin{array}{ll} \overline{X}_n = 1.42 & S^2 = 0.0054 \text{ I campione} \\ \overline{X}_n = 1.47 & S^2 = 0.0031 \text{ I campione} \end{array}$$

mentre per il quantile di Student usiamo

$$t_{0.05}(9) = 1.833 \qquad t_{0.025}(9) = 2.262 \qquad t_{0.005}(9) = 3.250 \qquad *$$

Gli intervalli sono

$(1.39, 1.45)$	per il I campione	e	$(1.46, 1.50)$	per il II campione	al 90 %
$(1.38, 1.46)$	per il I campione	e	$(1.45, 1.50)$	per il II campione	al 95 %
$(1.36, 1.47)$	per il I campione	e	$(1.44, 1.51)$	per il II campione	al 99 %

Gli intervalli al 90% non si sovrappongono.

* Usando i quantili della normale:

$$z_{90} = 1.645 \qquad z_{95} = 1.96 \qquad z_{99} = 2.576$$

	I campione	II campione
90% :	$(1.39, 1.45)$	$(1.46, 1.49)$
95% :	$(1.39, 1.45)$ #	$(1.45, 1.49)$
99% :	$(1.37, 1.46)$	$(1.45, 1.50)$

\# uguali a causa degli arrotondamenti

4. Il consumo giornaliero di gas per riscaldamento (per famiglia) in una città di latitudine media è una variabile casuale X distribuita secondo una legge normale di media $\mu = 2$ metri cubi standard (Smc) e deviazione standard $\sigma = 0.5$ Smc (numeri consistenti con i dati ISTAT). Qual è la probabilità che

   (i) in esattamente 2 giorni nel mese il consumo superi i 3 Smc giornalieri?

   (ii) che in tutto il mese di gennaio si consumino almeno 70 Smc di gas?

Soluzione. $X \sim N(2, 0.5)$

(i) $p = P(X > 3) = P\left(Z > \frac{3-\mu}{\sigma}\right) = P(Z > 2) =$

$= 1 - P(Z \leq 2) = 1 - \Phi(2) = 1 - 0.9772 = 0.0228$

$Y = \#$ di giorni in cui $X > 3$

$Y \sim B(30, p)$

$P(Y=2) = \binom{30}{2} p^2 (1-p)^{28} = \frac{30!}{2!\,28!} p^2 (1-p)^{28} =$

$= \frac{30 \cdot 29}{2} p^2 (1-p)^{28} = 15 \cdot 29\; p^2 (1-p)^{28} = 0.12$

(0.118548)

(ii) X_k consumo nel giorno k

$$P\left(\sum_{k=1}^{31} X_k \geq 70\right) = \qquad (n=31)$$

$$P\left(\frac{\sum_k X_k - n \cdot 2}{\sigma\sqrt{n}} \geq \frac{70-62}{0.5\sqrt{31}}\right) =$$

$$= P(Z \geq 2.87) = 1 - \Phi(2.87) =$$

$$= 1 - 0.9979 = 0.0021$$