PROVE D'ESAME DI CPS

28/05/2009

(1) Consideriamo 4 scatole contenenti 2 palline ciascuna indistinguibili al tatto. La scatola i ha una pallina numerata con 1 e l'altra numerata con i (quindi nella prima entrambe le palline sono numerate 1). Si estrae una pallina per scatola. Si denoti con X la variabile somma delle palline estratte e con Y la variabile numero di palline estratte con il numero 1. Si determini
  (a) la legge, la media e la varianza di X;
  (b) la probabilità di aver estratto la biglia con il 4 sapendo che $X = 7$;
  (c) la legge e la media di Y;
  (d) la covarianza di (X, Y).

(2) Sia

$$f(t) = \begin{cases} kx^2 e^{-\frac{x}{2}} & \text{se } x \geq 0; \\ 0 & \text{altrimenti.} \end{cases}$$

  (a) Si verifichi che per $k = \frac{1}{16}$ la funzione $f(t)$ è la densità di una variabile X;
  (b) Determinare la densità delle variabili $2X$ e $X + Y$, dove Y è ua variabile indipendente da X avente la stessa densità di X.

(3) Il nostro produttore di fiducia di albicocche produce albicocche aventi peso medio 98 grammi e scarto quadratico medio 5 grammi.
  (a) Qual è la probabilità che una cassa contenente 100 albicocche pesi più di 10 kg?
  Un giorno compriamo una cassa di albicocche di 10 kg da un altro produttore che afferma che le sue albicocche pesano mediamente 115 grammi. Andiamo a casa e contiamo le albicocche nella cassa: sono 98.
  (b) Era legittima l'affermazione del secondo produttore? (Si assuma uno scarto quadratico medio sempre pari a 5 grammi).

03/06/2009

(1) In un gioco di carte si pescano 3 carte a caso da un mazzo di 40 carte italiane. Si fanno 8 punti se le tre carte sono uguali e si fanno 2 punti se la somma delle carte è minore di 7 (e non sono tutte e tre uguali). Sia X la variabile aleatoria "punti fatti in una partita"
  (a) determinare la legge di X;
  (b) determinare la media e la varianza di X;
  (c) determinare la probabilità di fare almeno 18 punti in 117 partite.

Soluzione. COnsideriamo come spazio di probabilità Ω l'insieme di tutte le terne (non ordinate) di carte che possiamo pescare. Abbiamo $\binom{40}{3}$ possibili risultati. Per ottenere 2 punti dobbiamo avere somma delle carte minore di 7 senza avere 3 carte uguali. Questo accade in 5 modi possibili: 112,113,114,221,123. I primi 4 sono chiaramente equiprobabili, mentre l'ultimo ha una probabilità diversa. Si ha

$$P\{112\} = P\{113\} = P\{114\} = P\{221\} = \frac{\binom{4}{2}\binom{4}{1}}{\binom{40}{3}}$$

$$P\{123\} = \frac{\binom{4}{1}\binom{4}{1}\binom{4}{1}}{\binom{40}{3}}.$$

Si ha quindi $P\{X = 2\} = P\{112\} + P\{113\} + P\{114\} + P\{221\} + P\{123\} = 4\frac{\binom{4}{2}\binom{4}{1}}{\binom{40}{3}} + \frac{\binom{4}{1}\binom{4}{1}\binom{4}{1}}{\binom{40}{3}} = \frac{4}{247}$.

La probabilità di avere 3 carte uguali è 10 volte la probabilità di avere 3 assi. Quest'ultima è pari a $\frac{\binom{4}{3}}{\binom{40}{3}}$. Si ha quindi

$$P\{X = 8\} = 10P\{111\} = 10\frac{\binom{4}{3}}{\binom{40}{3}} = \frac{1}{247}.$$

Concludiamo che la legge di X è

$$p_X(k) = \begin{cases} \frac{1}{247} & \text{se } k = 8; \\ \frac{4}{247} & \text{se } k = 2; \\ \frac{242}{247} & \text{se } k = 0. \end{cases}$$

La media di X è $E[X] = 8\frac{1}{247} + 2\frac{4}{247} = \frac{16}{247}$. Per calcolare la varianza determiniamo la media di X^2, $E[X^2] = 64\frac{1}{247} + 4\frac{4}{247} = \frac{80}{247}$. Si ha quindi $Var(X) = E[X^2] - E[X]^2 = \frac{80}{247} - \frac{256}{(247)^2} = \frac{19504}{(247)^2}$.

Per il teorema limite centrale la somma Y dei punteggi ottenuti in 247 partite è approssimativamente normale con media $\mu_Y = 247\mu_X$ e varianza $Var(Y) = 247Var(X)$. Si ha quindi $\mu_Y = 16$ e $Var(Y) = \frac{19504}{247}$ quindi $\sigma_Y = \sqrt{Var(Y)} = 8,89$. Si ha quindi $P\{Y \geq 18\} = P\{\zeta_0 > \frac{17.5-16}{8.89}\} = P\{\zeta_0 > 0.17\} = 43.25\%$, dove ζ_0 è una variabile normale standard.

(2) Donatella ha 2 ristoranti che sfruttano una cucina unica. Siano X e Y le variabili numero di clienti per sera dei 2 ristoranti. Si sa che che X e Y seguono una legge di Poisson di parametri 20 e 30 rispettivamente.
  (a) Si determini la legge del numero totale di clienti;
  (b) Sapendo di aver avuto una sera complessivamente 45 clienti si determini la probabilità che il primo ristorante abbia avuto solo 10 clienti;
  (c) Verificare che la legge condizionale di X sapendo che $X + Y = 45$ è $B(45, 2/5)$.

Soluzione. La somma di 2 variabili di Poisson indipendenti è ancora una variabile di POisson avente per parametro la somma dei parametri. $X+Y$ è quindi di Poisson di parametro 50.

$$\begin{aligned} P\{X=10|X+Y=45\} &= \frac{P\{X=10, X+Y=45\}}{P\{X+Y=45\}} = \frac{P\{X=10, Y=35\}}{P\{X+Y=45\}} \\ &= \frac{P\{X=10\}P\{Y=35\}}{P\{X+Y=45\}} = \frac{e^{-20}\frac{20^{10}}{10!}e^{-30}\frac{30^{35}}{35!}}{e^{-50}\frac{50^{45}}{45!}} \\ &= \frac{\frac{20^{10}}{10!}\frac{30^{35}}{35!}}{\frac{50^{45}}{45!}} = \binom{45}{10}(2/5)^{10}(3/5)^{35} = 0,6\% \end{aligned}$$

Il punto (c) si ottiene come dal (b) scrivendo k al posto di 10 (e $45-k$ al posto di 35).

(3) In un sondaggio per un referendum 55 persone hanno dichiarato di voler votare SI e 45 persone hanno dichiarato di voler votare NO.

(a) Sottoporre a test l'ipotesi in cui il SI vince al referendum.

Al referendum effettivamente vince il SI con il 53% dei voti.

(b) Qual è la probabilità che in un campione scelto a caso di 100 votanti almeno la metà abbia votato SI?

22/06/2009

(1) Si considerino tre scatole contenenti 10 palline ciascuna. La prima ne ha 9 rosse e una blu, la seconda ne ha 9 blu e una gialla, la terza ne ha 9 gialle e una rossa. Pescando una pallina per ogni scatola
   (a) qual è la probabilità di pescare 3 palline di colore differente?
   (b) Sapendo di aver pescato 3 palline di colore differente, qual'è la probabilità di aver pescato la pallina rossa dalla prima scatola?
   (c) determinare la legge e la media della variabile X che conta il numero di palline blu estratte.

Soluzione. (a) Sia RBG l'evento estrazione, nell'ordine, di una rossa, una blu e una gialla. Similmente si definisca BGR. La probabilit1'a richiesta è quindi

$$P(RBG \cup BGR) = P(RBG) + P(BGR) = (0.9)^3 + (0.1)^3 = 73\%.$$

(b) Detto R_1 l'evento "estrazione di una rossa dalla prima scatola" dobbiamo calcolare

$$P(R_1|RBG \cup BGR) = \frac{P(R_1 \cap (RBG \cup BGR)}{P(RBG \cup BGR} = \frac{P(RBG)}{P(RBG \cup BGR} = \frac{0.729}{0.73} = 99.9\%$$

(c) Sia X_1 la variabile che vale 1 se la prima estratta è blu e 0 altrimenti. Si definisca similmente X_2. Si ha quindi $X = X_1 + X_2$. Si ha $X_1 \sim B(1, \frac{1}{10})$ e $X_2 \sim B(1, \frac{9}{10})$ e pertanto $E[X] = E[X_1] + E[X_2] = 1/10 + 9/10 = 1$.
La variabile X può assumere solo i valori $0, 1, 2$. Determiniamone la legge

$$\begin{aligned}
P(X = 0) &= P(X_1 = 0, X_2 = 0) = P(X_1 = 0)P(X_2 = 0) = 9/10 \cdot 1/10 = 9\% \\
P(X = 1) &= P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0) = (9/10)^2 + (1/10)^2 = 82\% \\
P(X = 2) &= P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) = 1/10 \cdot 9/10 = 9\%
\end{aligned}$$

(2) Si consideri la seguente funzione

$$F(t) = \begin{cases} 0 & \text{se } t < 0 \\ -t^2 + 2t & \text{se } 0 \le t < 1 \\ 1 & \text{altrimenti} \end{cases}$$

   (a) Verificare che $F(t)$ è la funzione di ripartizione di una variabile aleatoria continua X.
   (b) Deteminare la legge di X.
   (c) Determinare la media di X.

Soluzione. (a) La funzione $F(t)$ è continua, non decrescente ed ha i limiti corretti quindi è una funzione di ripartizione.
(b) La densità si calcola per derivazione della funzione di ripartizione ed è quindi

$$f(t) = \begin{cases} 0 & \textit{se } t < 0 \textit{ o } t > 1 \\ -2t + 2 & \textit{altrimenti} \end{cases}$$

(c)

$$E[X] = \int_{-\infty}^{+\infty} t f(t)\, dt = \int_0^1 (-2t^2 + 2t) dt = [-\frac{2}{3}t^3 + t^2]_0^1 = \frac{1}{3}$$

(3) Per sottoporre a test l'effetto di un nuovo fertilizzante sulla produzione di frumento un appezzamento viene diviso in 60 parcelle di uguale superficie. Il nuovo fertilizzante è stato usato in 30 parcelle e il vecchio nelle restanti. Nelle parcelle in cui si è utilizzato il nuovo fertilizzante si sono prodotti 18,2 quintali con uno scarto quadratico medio di 0,63, mentre nelle altre si sono prodotti 17,8 quintali con uno scarto qudartico medio di 0,54. Si sottoponga a test l'ipotesi in cui il nuovo fertilizzante è migliore del vecchio al livello di significatività del 5% e dell' 1%.

Soluzione. Siano μ_0 *e* μ_1 *la quantità media di frumento prodotta rispettivamente con il vecchio e con il nuovo fertilizzante in una parcella. Formuliamo l'ipotesi nulla* $H_0 : \mu_0 = \mu_1$ *e l'ipotesi alternativa* $\mu_0 < \mu_1$. *Sotto l'ipotesi nulla la distribuzione delle differenze campionarie ha media nulla e scarto quadratico medio pari a*

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 0,15$$

La variabile standardizzata relativa ai nostri campioni è quindi $\frac{18,2-17,8}{0,15} = 2,67$. *Rifiutiamo pertanto l'ipotesi sia a livello di significatività del 5% che dell'1% concludendo che il nuovo fertilizzante è senz'altro più efficace.*

13/07/2009

(1) Si hanno a disposizione gelati in 6 gusti. I primi 3 gusti sono scelti mediamente da 1 persona su 9 e gli altri gusti da 2 persone su 9. Offriamo un gelato ciascuno ai nostri 10 amici. Si consideri la seguente variabile aleatoria

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo gusto è scelto da qualcuno} \\ 0 & \text{altrimenti.} \end{cases}$$

(a) Qual'è la legge delle X_i?
(b) Sono indipendenti?
(c) Quanti gusti mi devo aspettare di utilizzare (mediamente)?

(2) Si consideri la seguente funzione

$$F(t) = \begin{cases} 0 & \text{se } t < 0 \\ -2t^3 + 3t^2 & \text{se } 0 \leq t < 1 \\ 1 & \text{altrimenti} \end{cases}$$

(a) Verificare che $F(t)$ è la funzione di ripartizione di una variabile aleatoria continua X.
(b) Deteminare la densità di X.
(c) Determinare la media di X.

(3) Si vuole sottoporre a test l'ipotesi in cui le persone abitanti nella città A mangiano di più di quelle abitanti nella città B. Si conduce un indagine sul consumo calorico medio giornaliero nelle due città e si ottengono rispettivamente i seguenti risultati

Kcal	Persone
3000	10
2500	12
2000	18
1500	10

Kcal	Persone
3000	11
2500	15
2000	17
1500	12
1000	5

Cosa possiamo concludere?

10/09/2009

(1) Avendo a disposizione l'elenco del telefono decidiamo di fare degli scherzi telefonici agli abitanti dei paesi A,B,C. Il paese A ha circa il doppio degli abitanti di B e C. Si sceglie un numero a caso e:
- Se il numero è del paese A facciamo lo scherzo 1;
- Se il numero è del paese B facciamo lo scherzo 2;
- Se il numero è del paese C facciamo indifferentemente lo scherzo 1 o lo scherzo 2.

Facendo complessivamente 50 scherzi sia X il numero di volte in cui abbiamo fatto lo scherzo 1 ed N_A il numero di volte che abbiamo chiamato un abitante del paese A.

(a) Determinare la probabilità che il primo scherzo effettuato sia lo scherzo 1;
(b) Determinare la legge di X;
(c) Determinare la legge di N_A;
(d) Determinare la probabiltià di aver chiamato sempre abitanti del paese A sapendo di aver sempre fatto lo scherzo 1.

(a) Sia U l'evento scherzo 1 alla prima chiamata, A, B, C gli eventi prima chiamata verso il paese A, B, $C >$ rispettivamente. Allora

$$\begin{aligned} P(U) &= P(U \cap A) + P(U \cap B) + P(U \cap C) \\ &= P(A)P(U|A) + P(B)P(U|B) + P(C)P(U|C) \\ &= \frac{1}{2}1 + \frac{1}{4}0 + \frac{1}{4}\frac{1}{2} = \frac{5}{8}. \end{aligned}$$

(b) La variabile X è una variabile $B(50, \frac{5}{8})$ e quindi la legge di X è

$$P(X = k) = \binom{50}{k}(5/8)^k(3/8)^{50-k};$$

(c) La variabile N_A è una variabile $B(50, \frac{1}{2}$;
(d) Si ha

$$\begin{aligned} P(N_A = 50|X = 50) &= \frac{P(N_A = 50)P(X = 50|N_A = 50)}{P(X = 50)} \\ &= \frac{(1/2)^{50}1}{(5/8)^{50}} = (4/5)^{50} = 0,0014272\% \end{aligned}$$

Questo evento è quindi praticamente impossibile.

(2) Un componente elettronico è formato da 2 elementi in parallelo collegati in serie ad un terzo elemento. I 3 elementi sono uguali ed aventi tempo di vita esponenziale di media 1 anno e sono indipendenti tra loro. Sia T la variabile tempo di vita del componente (si ricorda che il componente rimane funzionante finchè funziona almeno uno dei primi 2 elementi e il terzo).

(a) Determinare la funzione di ripartizione di T;
(b) Determinare la densità di T;
(c) Deteminare la media di T.

(a)Siano T_i, $i = 1, 2, 3$ le variabili tempo di vita dei 3 elementi. Per ipotesi queste sono tutte variabili esponenziali di parametro $\lambda = 1$. (Ricordo che la media di una variabile esponenziale è l'inverso del parametro λ.)

Quindi $F_1(t) = F_2(t) = F_3(t) = 1-e^{-t}$, dove F_i indica la funzione di ripartizione della variabile T_i. La variabile T è data da $T = \min(\max(T_1, T_2), T_3)$. Sia G la funzione di ripartizione di $\max(T_1, T_2)$. Si ha

$$\begin{aligned} G(t) &= P(\max(T_1, T_2)) \le t) = P(T_1 \le t \text{ e } T_2 \le t) \\ &= P(T_1 \le t)P(T_2 \le t) \\ &= (1 - e^{-t})^2 \end{aligned}$$

per l'indipendenza delle variabili T_1 e T_2. Sia ora F la funzione di ripartizione di T. Abbiamo

$$\begin{aligned} 1 - F(t) &= P(T > t) = P(\min(\max(T_1, T_2), T_3) > t) \\ &= P(\max(T_1, T_2) > t)P(T_3 > t) \\ &= (1 - G(t))(1 - F_3(t)) = (2e^{-t} - e^{-2t})e^{-t} \\ &= 2e^{-2t} - e^{-3t}, \end{aligned}$$

e quindi

$$F(t) = 1 - 2e^{-2t} + e^{-3t}.$$

(b) $f(t) = F'(t) = 4e^{-2t} - 3e^{-3t}$.
(c) $E[T] = \int tf(t)\,dt = \int(4te^{-2t} - 3te^{-3t})dt = 2\int 2te^{-2t}dt - \int 3te^{-3t}dt = 2\frac{1}{2} - \frac{1}{3} = 2/3$. Abbiamo sfruttato l'integrale

$$\int \lambda te^{\lambda t}dt = \frac{1}{\lambda}.$$

La vita media del componente è quindi di 2/3 di anno, cioè 8 mesi.

(3) Si conduce un'indagine e su 100 bambini di campagna ne risultano 57 senza problemi di allergia, mentre su 1000 bambini di città ne risultano 548 senza problemi di allergia.
(a) Determinare gli intervalli di confidenza al 5% e all 1% per la differenza tra le 2 proporzioni campionarie in esame.
(b) Determinare con quale significatività si può accettare l'ipotesi in cui i bambini di campagna abbiano mediamente meno problemi di allergia dei bambini di città.

(a) Siano p_A e p_B le proporzioni di bambini ci campagna e città rispettivamente che non hanno problemi di allergia. La variabile differenza di proporzioni campionarie $P_A - P_B$ si distribuisce normalmente con media $p_A - p_B$ e scarto quadratico medio

$$\sigma_{P_A - P_B} = \sqrt{\sigma^2_{P_A} + \sigma^2_{P_B}} = \sqrt{p_A q_A/N_A + p_B q_B/N_B)}.$$

Approssimiamo questa quantità utilizzando i valori registrati

$$\sigma_{P_A - P_B} = \sqrt{P_A(1 - P_A)/N_A + P_B(1 - P_B)/N_B} = \sqrt{0,002451 + 0,0002476} = 0,052.$$

Concludiamo che la differenza $p_A - p_B$ è compresa nell'intervallo $[P_A - P_B - 1,96\sigma_{P_A - P_B}, P_A - P_B + 1,96\sigma_{P_A - P_B}] = [-0,08; 0,12]$ con significatività al 95% e similmente il caso al 99% sostituendo 1,96 con 2,58.

(b) Facciamo l'ipotesi nulla $H_0 : p_A - p_B = 0$ e l'ipotesi alternativa $H_1 : p_A - p_B > 0$. Assumendo l'ipotesi nulla poniamo $p = p_A = p_B$ dove p è la media pesata dei valori registrati $p = \frac{N_A P_A + N_B P_B}{N_a + N_B} = \frac{100 \cdot 0,57 + 1000 \cdot 0,548}{1100} = 0,55$. Il valore registrato in unità standard vale $z = \frac{P_A - P_B}{\sqrt{p(1-p)(\frac{1}{N_A} + \frac{1}{N_B})}} = \frac{0,022}{\sqrt{0,55 \cdot 0,45 \frac{11}{1000}}} = 0,42$. Tale valore è decisamente troppo basso per rifiutare l'ipotesi nulla e quindi concludiamo che la differenza registrata è solo dovuta al caso.

18/01/2010

(1) In un gioco del gratta e vinci si ha probabilità 30% di vincere 5 Euro, il 20% di vincere 10 Euro e il 50% di non vincere nulla. Un nostro amico compra 5 biglietti.
(a) Determinare la speranza matematica.
(b) Determinare la probabilità di vincere qualcosa.
(c) Determinare la probabilità di trovare almeno un biglietto vincente da 10 Euro.
(d) Il nostro amico va alla cassa per ritirare quello che ha vinto. Qual è a questo punto la probabilità di aver trovato un biglietto da 10 Euro?

(2) Viene esaminato il diametro di un campione casuale di 25 mele di un certo frutteto. Si ottengono i seguenti risultati.

Diametro in cm	Quantità
6-7	3
7-8	4
8-9	7
9-10	6
10-11	5

(a) Determinare la media campionaria $\bar{X}$ e lo scarto quadratico medio campionario s.
(b) Determinare l'intervallo di confidenza al 95% per la media dei diametri di tutte le mele prodotte nel frutteto.
(c) Una mela va in categoria I se il suo diametro è di almeno 8 cm. Determinare l'intervallo di confidenza al 95% per la proporzione di mele prodotte nel frutteto in categoria I.

15/02/2010

(1) In una scatola ci sono 5 palline rosse e 5 blu. Le palline vengono pescate una alla volta. Siano X e Y le variabili "prima pescata di una pallina rossa" nel caso rispettivamente di reinserimento della pallina pescata e di non reinserimento della pallina pescata.
   (a) Determinare, la legge, la media e la varianza delle variabili X ed Y.
   (b) Commentare l'affermazione " le variabili X ed Y sono indipendenti".

(2) Si consideri la popolazione Ω data dagli studenti del nostro corso di laurea. Si consideri la variabile X su Ω data da il numero di contatti su Facebook. Si supponga che la media di X sia 50 con uno scarto quadratico medio di 15.
   (a) Si determini la probabilità che 50 studenti scelti a caso abbiano complessivamente almeno 3000 contatti.
   (b) Quanto dovrebbe essere la media di X (supponendo di mantenere lo stesso scarto quadratico medio) per fare in modo che la probabiltià che 50 studenti scelti a caso abbiano complessivamente almeno 3000 contatti sia del 90%?

01/06/2010

(1) (14pt) Olimpiadi, finale dei 100m maschili, 8 finalisti. Si sa che i 4 atleti nelle corsie centrali hanno probabilità $\frac{2}{3}$ di correre in meno di 10 secondi. I 4 atleti delle corsie laterali hanno probabilità $\frac{1}{3}$ di correre in meno di 10 secondi.
Si considerino le seguenti variabili:

$$\begin{aligned} X &= \#\{\text{atleti che corrono sotto i 10 secondi}\} \\ Y &= \begin{cases} 1 & \text{se almeno 6 atleti corrono sotto i 10 secondi} \\ 0 & \text{altrimenti} \end{cases} \end{aligned}$$

(a) Determinare la media e la varianza di X;
Sia
$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esimo atleta corre in meno di 10 secondi} \\ 0 & \text{altrimenti.} \end{cases}$$
Le X_i sono tutte variabili di Bernoulli. Per $i = 1, 2, 7, 8$ si ha $X_i \sim B(1, \frac{1}{3})$ e per $i = 3, 4, 5, 6$ si ha $X_i \sim B(1, \frac{2}{3})$. Si ha $X = X_1 + X_2 + \cdots + X_8$ e quindi
$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_8] = 4 \cdot \frac{1}{3} + 4 \cdot \frac{2}{3} = 4.$$
Inoltre, siccome le X_i si possono assumere indipendenti ed hanno tutte varianza $\frac{1}{3} \cdot \frac{2}{3}$ si ha
$$Var(X) = Var(X_1) + Var(X_2) + \cdots + Var(X_8) = 8 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{16}{9}.$$

(b) Determinare la densità di Y;
Siccome Y può assumere i soli valori 0 e 1, Y è una variabili di Bernoulli $B(1, p)$. Per determinare il parametro p utilizziamo le variabili ausiliarie $X' = X_1+X_2+X_7+X_8$ e $X'' = X_3+X_4+X_5+X_6$. Si ha chiaramente $X' \sim B(4, \frac{1}{3})$ e $X'' \sim B(4, \frac{2}{3})$. calcoliamo
$$p = P(Y = 1) \quad = \quad P(X = 6) + P(X = 7) + P(X = 8).$$
Ma
$$\begin{aligned} P(X = 6) &= P(X' = 4)P(X'' = 2) + P(X' = 3)P(X'' = 3) + P(X' = 2)P(X'' = 4) \\ &= (\frac{1}{3})^4 \binom{4}{2}(\frac{1}{3})^2(\frac{2}{3})^2 + 4(\frac{1}{3})^3(\frac{2}{3})\ 4(\frac{2}{3})^3(\frac{1}{3}) + (\frac{1}{3})^4 \binom{4}{2}(\frac{1}{3})^2(\frac{2}{3})^2\ (\frac{2}{3})^4 \\ &= \frac{24}{3^8} + \frac{256}{3^8} + \frac{384}{3^8} = \frac{664}{3^8}. \end{aligned}$$
Inoltre
$$\begin{aligned} P(X = 7) &= P(X' = 4)P(X'' = 3) + P(X' = 3)P(X'' = 4) \\ &= (\frac{1}{3})^4\ 4(\frac{2}{3})^3(\frac{1}{3}) + 4(\frac{1}{3})^3(\frac{2}{3})\ (\frac{2}{3})^4 \\ &= \frac{32}{3^8} + \frac{128}{3^8} = \frac{160}{3^8}. \end{aligned}$$
Infine
$$\begin{aligned} P(X = 8) &= P(X' = 4)P(X'' = 4) \\ &= (\frac{1}{3})^4\ (\frac{2}{3})^4 \\ &= \frac{16}{3^8} \end{aligned}$$

Concludendo
$$P(Y=1) = \frac{664}{3^8} + \frac{160}{3^8} + \frac{16}{3^8} = \frac{840}{3^8} = 0.128,$$
e quindi la densità di Y è $B(1, 0.128)$.

(c) Supponendo che almeno 6 atleti abbiano corso in meno di 10 secondi, qual è la probabilità che tutti i quattro atleti delle corsie centrali abbiano corso in meno di 10 secondi? La probabilità richiesta è
$$\begin{aligned} P(X''=4|Y=1) &= \frac{P(X''=4 \cap Y=1)}{P(Y=1)} \\ &= \frac{P(X''=4)\Big(P(X'=2)+P(X'=3)+P(X'=4)\Big)}{P(Y=1)} \\ &= \frac{(\frac{2}{3})^4\Big(\binom{4}{2}(\frac{1}{3})^2(\frac{2}{3})^2 + 4(\frac{1}{3})^3(\frac{2}{3}) + (\frac{1}{3})^4\Big)}{\frac{840}{3^8}} \\ &= \frac{22}{35} = 0.63. \end{aligned}$$

(d) Vi sembra conveniente una scommessa con quota 10:1 (cioè si vincono 10 euro per ogni euro puntato) sul fatto che almeno 6 atleti corrano in meno di 10 secondi?
Detta T la variabile T="soldi vinti dopo una scommessa di un euro" si ha:
$$E[T] = 10 \cdot P(Y=1) + 0 \cdot P(Y=0) = 10 \cdot 0.128 = 1.28.$$
Con una scommessa di 1 euro si vincono mediamente 1.28 euro e quindi la scommessa è conveniente.

(2) (8pt) Il tempo di attesa di eruzione dell'Etna ha densità esponenziale di media 5 anni.

(a) Si determini la probabilità di avere un'eruzione nei prossimi 2 anni.
La variabile T tempo di attesa per un'eruzione è esponenziale di parametro $\lambda = 1/5$. Quindi la funzione di ripartizione di T è data da:
$$F_T(t) = \begin{cases} 1 - e^{-t/5} & \text{se } t > 0 \\ 0 & \text{altrimenti.} \end{cases}$$
Si ha quindi
$$P(0 < T < 2) = F_T(2) - F_T(0) = 1 - e^{-2/5} = 0.33.$$

(b) Supponendo di non aver avuto eruzioni nei prossimi 2 anni, si determini la probabilità di avere un'eruzione nei successivi 2 anni.
Ricordando che una variabile esponenziale soddisfa la proprietà della mancanza di memoria si ha
$$P(2 < T < 4|T > 2) = P(0 < T < 2) = 1 - e^{-2/5} = 0.33.$$

(3) (8pt) Un esperimento consiste nello scegliere a caso 50 numeri reali compresi tra 1 e 3.
Sia X la variabile somma di questi 50 numeri.

(a) Determinare la media e la varianza di X.
Siano $X_1, X_2, \ldots, X_{50}$ i 50 numeri estratti. Le variabili X_i sono indipendenti e uniformi nell'intervallo $[1, 3]$. Hanno quindi tutte media 2 e varianza $\frac{(4-2)^2}{12} = \frac{1}{3}$. Ne segue che
$$E[X] = 50 \cdot E[X_1] = 50 \cdot 2 = 100,$$
e
$$Var(X) = 50 \cdot Var(X_1) = \frac{50}{3}.$$

(b) Determinare la probabilità che X sia compreso tra 95 e 110.
Per il teorema limite centrale possiamo approssimare X con una variabile normale $\zeta \sim N(100, \frac{50}{3})$.
Si ha

$$\begin{aligned}
P(95 < \zeta < 110) &= P\left(\frac{95-100}{\sqrt{\frac{50}{3}}} < \zeta_0 < \frac{110-100}{\sqrt{\frac{50}{3}}}\right) \\
&= P(-1.22 < \zeta_0 < 2.45) \\
&= \Phi(2.45) - (1 - \Phi(1.22)) \\
&= 0.99286 + 0.88877 - 1 \\
&= 0.88.
\end{aligned}$$

15/06/2010

(1) (12pt) In un sacchetto abbiamo 11 monete, numerate da 0 a 10, indistinguibili al tatto. La probabilità che la moneta numerata i dia testa è pari a $i/10$. Si prende una moneta a caso e si lancia 3 volte. Si considerino le variabili X_1, X_2, X_3 date da

$$X_j = \begin{cases} 1, & \text{se il lancio } j \text{ dà testa;} \\ 0, & \text{altrimenti.} \end{cases}$$

(a) Determinare le densità delle variabili X_1, X_2, X_3.
Le tre variabili hanno evidentemente la stessa densità, che sarà una $B(1, p)$. Per determinare p poniamo A_i l'evento "scelta della moneta numerata i" e utilizziamo la formula delle probabilità totali:

$$\begin{aligned} p &= P(X_1 = 1) \\ &= P(A_0)P(X_1 = 1|A_0) + P(A_1)P(X_1 = 1|A_1) + \cdots + P(A_{10})P(X_1 = 1|A_{10}) \\ &= \frac{1}{11} \cdot \frac{0}{10} + \frac{1}{11} \cdot \frac{1}{10} + \cdots + \frac{1}{11} \cdot \frac{10}{10} \\ &= \frac{1}{2}. \end{aligned}$$

Concludiamo che le variabili X_1, X_2, X_3 hanno tutte denistà $B(1, \frac{1}{2})$.

(b) Studiare l'indipendenza delle variabili X_1, X_2, X_3.
Se il primo lancio dà testa ci aspettiamo che sia più facile aver scelto una moneta con numero alto e che quindi la probabilità che anche il secondo lancio dia testa ci aspettiamo che aumenti, e che quindi le variabili siano dipendenti. Lo verifichiamo facendo il conto

$$\begin{aligned} P(X_1 = 1, X_2 = 1) &= P(A_0)P(X_1 = 1, X_2 = 1|A_0) + P(A_1)P(X_1 = 1, X_2 = 1|A_1) + \cdots \\ &= \frac{1}{11} \cdot \left(\frac{0}{10}\right)^2 + \frac{1}{11} \cdot \left(\frac{1}{10}\right)^2 + \cdots + \frac{1}{11} \cdot \left(\frac{10}{10}\right)^2 \\ &= \frac{1}{11 \cdot 100}(0 + 1 + 4 + 9 + 16 + \cdots + 100) \\ &= \frac{385}{1100} = \frac{7}{20} = 35\% \end{aligned}$$

Si ha quindi che $P(X_1 = 1, X_2 = 1) \neq P(X_1 = 1)P(X_2 = 1) = \frac{1}{4} = 25\%$ e quindi le variabili X_1 e X_2 sono dipendenti.

(c) Determinare l'indice di correlazione ρ_{X_1,X_2}.
Calcoliamo intanto la covarianza $\sigma_{X_1,X_2} = E[X_1X_2] - E[X_1]E[X_2]$. Per calcolare $E[X_1X_2]$ osserviamo che X_1X_2 può assumere solo i valori 0 e 1 e quindi è di Bernoulli e concludiamo che

$$E[X_1X_2] = P(X_1X_2 = 1) = P(X_1 = 1, X_2 = 1) = \frac{7}{20}.$$

Si ha quindi

$$\rho_{X_1,X_2} = \frac{\sigma_{X_1,X_2}}{\sigma_{X_1}\sigma_{X_2}} = \frac{\frac{7}{20} - \frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2}} = \frac{2}{5}.$$

(d) Supponendo di aver ottenuto 3 teste nei 3 lanci, qual è la probabilità di aver estratto la moneta numerata con il 9 o la moneta numerata con il 10?

Sia T l'evento "3 teste nei 3 lanci". Si ha

$$\begin{aligned} P(T) &= P(A_0)P(T|A_0) + P(A_1)P(T|A_1) + \cdots + P(A_{10})P(T|A_{10}) \\ &= \frac{1}{11 \cdot 1000}(0^3 + 1^3 + 2^3 + \cdots + 10^3) \\ &= \frac{3025}{11000} = \frac{11}{40} = 27.5\%. \end{aligned}$$

Per la formula di Bayes si ha

$$P(A_9|T) = \frac{P(T|A_9)P(A_9)}{P(T)} = \frac{\left(\frac{9}{10}\right)^3 \cdot \frac{1}{11}}{\frac{11}{40}} = \frac{729}{3025},$$

e

$$P(A_{10}|T) = \frac{P(T|A_{10})P(A_{10})}{P(T)} = \frac{1 \cdot \frac{1}{11}}{\frac{11}{40}} = \frac{40}{121}.$$

La probabilità richiesta è dunque

$$\frac{729}{3025} + \frac{40}{121} = \frac{1729}{3025} = 57\%.$$

(2) (10pt) Si consideri la seguente funzione $f : \mathbb{R} \to \mathbb{R}$

$$f(t) = \begin{cases} e^{-2t}, & \text{se } t \geq 0; \\ e^{2t}, & \text{se } t \leq 0. \end{cases}$$

(a) Verificare che f è la densità di una variabile aleatoria continua X;
Si ha $f(t) \geq 0$ e

$$\int_{-\infty}^{+\infty} = 2\int_0^{+\infty} e^{-2t}dt = 1,$$

dove si è sfruttato che la funzione f è pari, e quindi f è una densità continua.

(b) Determinare la funzione di ripartizione di X;
Se $t < 0$ si ha

$$F(t) = \int_{-\infty}^{t} e^{2s}ds = \left[\frac{1}{2}e^{2s}\right]_{-\infty}^{t} = \frac{1}{2}e^{2t}.$$

Se $t > 0$ si ha

$$F(t) = \int_{-\infty}^{t} f(s)ds = \int_{-\infty}^{0} e^{2s}ds + \int_0^t e^{-2s}ds = \frac{1}{2} + \left[-\frac{1}{2}e^{-2s}\right]_0^t = 1 - \frac{1}{2}e^{-2t}.$$

(c) Determinare la media di X;
Si ha $E[X] = 0$ perché la densità è pari.

(d) Determinare la probabilità che $X > 1$;

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F_X(1) = 1 - (1 - \frac{1}{2}e^{-2}) = \frac{1}{2}e^{-2} = 6.7\%$$

(*) Determinare la varianza di X.

(3) (8pt) Due amici Arturo e Basilio giocano a testa e croce, lanciando ciascuno una moneta 50 volte. Arturo vince se nei suoi lanci testa si presenterà almeno 5 volte in pià che nei lanci di Basilio.

(a) Determinare la probabilità che Arturo vinca (supponendo che le monete non siano truccate).
Dette A e B il numero di teste ottenute da Arturo e Basilio e $\bar{A}_{50}$ e $\bar{B}_{50}$ il numero medio di teste ottenute da Arturo e Basilio, la probabilità richiesta è (con la correzione di continuità)

$$P(A - B \geq 4.5) = P(\frac{A}{50} - \frac{B}{50} \geq \frac{4.5}{50}) = P(\bar{A}_{50} - \bar{B}_{50} \geq 0.09).$$

Approssimiamo $\bar{A}_{50} - \bar{B}_{50}$ con una variabile normale di media $E[A] - E[B] = 0$ e varianza $Var(\bar{A}_{50}) + Var(\bar{B}_{50}) = \frac{\frac{1}{2}\cdot\frac{1}{2}}{50} + \frac{\frac{1}{2}\cdot\frac{1}{2}}{50} = \frac{1}{100} = 0.01$. Si ha quindi

$$P(\bar{A}_{50} - \bar{B}_{50} \geq 0.1) = P(\zeta_0 > \frac{0.09 - 0}{\sqrt{0.01}}) = P(\zeta_0 > 0.9) = 18.4\%$$

(b) Se nei lanci di Arturo testa si è presentata 30 volte determinare l'intervallo di confidenza al 95% per la probabilità che la moneta di Arturo dia testa in un singolo lancio.
L'intervallo di confidenza ha come estremi

$$\frac{30}{50} \pm z_{95}\sqrt{\frac{\frac{3}{5}\cdot\frac{2}{5}}{50}} = 0.60 \pm 1.96 \cdot 0.07 = 0.60 \pm 0.14.$$

L'intervallo di confidenza è pertanto $[0.46, 0.74]$.

13/07/2010

(1) Negli ovetti di cioccolato con sorpresa si possono trovare 6 tipi di sorprese. Cinque di queste, numerate da 1 a 5, hanno probabilità $\frac{11}{60}$ ciascuna di essere trovate dentro un ovetto; l'ultima, più rara e numerata con 6, ha probabilità $\frac{1}{12}$. Si comprano 10 ovetti e poniamo

$$X_i = \begin{cases} 1, & \text{se abbiamo trovato almeno una volta la sorpresa } i; \\ 0, & \text{se non abbiamo trovato la sorpresa } i. \end{cases}$$

(a) Determinare la densità delle variabili X_i.
(b) Quante sorprese diverse ci dobbiamo aspettare di trovare complessivamente?
(c) Determinare la probabilità di trovare la sorpresa 1, ma non la sorpresa 2.
(d) (Extra) Come affronteresti il seguente problema: quanti ovetti devo comprare, mediamente, per poter collezionare tutte e 6 le sorprese?

(2) Il giardino di casa è illuminato da 2 lampioni. Il primo ha una lampadina di tipo A ed il secondo 2 lampadine di tipo B. Entrambi i tipi di lampadine hanno tempo di vita medio di densità esponenziale, il tipo A di media 4 anni, il tipo B di media 3 anni.

(a) Si determini la densità della variabile "tempo di vita del secondo lampione".
(b) Quale dei 2 lampioni mediamente illuminerà più a lungo?
(c) Determinare, mediamente, per quanto tempo rimarranno accese tutte e tre le lampadine.

(3) La densità congiunta $p(x, y)$ di 2 variabili aleatorie X ed Y è riassunta nella seguente tabella a doppia entrata

$X \backslash Y$	-1	0	1
1	.05	0	.05
4	.15	.30	0
5	0	.25	.20

Ad esempio si ha $p(4, 0) = 0.30$.

(a) Determinare le densità marginali di X ed Y.
(b) Le variabili X ed Y sono indipendenti?
(c) Determinare il valore atteso $E[X + Y]$.

28/09/2010

(1) Si considerino due quantità A ed B che ogni giorno possono aumentare o diminuire di 1 in modo equamente probabile ed indipendentemente l'una dall'altra. Ad esempio, se oggi A vale 4, domani ha il 50% di probabilitÃ di valere 3 e il 50% di valere 5. Si supponga che i valori iniziali di A e B siano rispettivamente 4 e 8. Siano X ed Y le variabili aleatorie date dai valori assunti da A ed B dopo 4 giorni.
   (a) Determinare la densità delle variabili X ed Y. Che tipo di variabili sono?
   (b) Detta Z la variabile che vale 1 se $X = Y$ e 0 altrimenti, si determini la densità della variabile Z.
   (c) Si discuta l'indipendenza delle variabili X e Z.

(2) Una certa pianta ha tempo di vita esponenziale di media 8 anni. Compro due di queste piante. Per quanto tempo, mediamente, vivrà almeno una delle due piante?

(3) In un campione casuale di 300 panettoni di una marca A si è registrata una quantità media di uvetta pari a 50 grammi con uno scarto quadratico medio di 2 grammi, mentre in un campione casuale di 200 panettoni di una marca B si è registrata una quantità media di uvetta pari a 49 grammi e uno scarto quadratico medio pari a 2 grammi.
   (a) Determinare i limiti di confidenza al 95% per la differenza delle medie delle quantità di uvetta presenti nei panettoni delle due marche.
   (b) Con quale confidenza possiamo affermare che la quantità media di uvetta nei panettoni della marca A è superiore alla quantità media di uvetta nei panettoni della marca B?

02/02/2011

(1) Si lancia un dado per 4 volte. Sia X la variabile che conta quante volte è uscito 6 e Y la variabile che conta il numero di volte in cui è uscito un numero dispari seguito da un numero pari.
   (a) Determinare le densità marginali delle variabili X ed Y.
   (b) Detta $p_{X,Y}$ la densità congiunta di X ed Y determinare $p_{X,Y}(2,1)$ e $p_{X,Y}(3,1)$.
   (c) Calcolare $P(X < 1|Y < 1)$.
   (d) Le variabili X ed Y sono indipendenti? Determinare la covarianza $\mathrm{Cov}(X,Y)$.

(2) Sia X una variabile continua uniforme nell'intervallo $[0,1]$ ed Y la variabile (discreta) data da $Y = \lfloor 5X \rfloor$, dove $\lfloor x \rfloor$ indica la parte intera di x, cioè il più grande numero intero minore o uguale a x. Determinare la densità di Y.

(3) Per determinare l'età media dei suoi abbonati, una piattaforma televisiva intervista un campione di 50 clienti riscontrando un valore medio di 36. Sapendo che l'età media degli abbonati si distribuisce normalmente e che $\sigma = 12$,
   (a) Determinare i limiti di confidenza al 95% per l'età media di tutti gli abbonati;
   (b) si supponga di voler ridurre l'ampiezza dell'intervallo di confidenza al 95%, in modo tale che gli estremi distino dal valore centrale dell'intervallo ± 2 anni. Quanto deve essere grande il campione?

06/06/2011

(1) Due dadi, uno rosso e uno blu, vengono lanciati simultaneamente per 3 volte. Si considerino le seguenti variabili aleatorie:
- X = Numero di volte in cui un dado ha dato come risultato 1 o 6;
- Y = Numero di lanci in cui il dado rosso ha dato risultato minore o uguale a 3;
- Z = Numero di lanci in cui la somma dei dadi ha dato 7.

(a) Determinare la densità delle variabili X, Y e Z;
(b) Stabilire se X ed Y sono indipendenti;
(c) Stabilire se Y e Z sono indipendenti;
(d) Stabilire se X e Z sono indipendenti.

(2) Sia X una variabile continua uniforme nell'intervallo $[0, 2]$. Si consideri la variabile $Y = \frac{1}{X+1}$.
(a) Determinare $P(Y < \frac{1}{2})$.
(b) Determinare la funzione di ripartizione di Y;
(c) Determinare la densità di Y.

(3) In un'indagine statistica su 100 studenti sulle ore di studio giornaliere dedicate alla preparazione dell'esame di CPS si sono ottenuti i seguenti risultati:

Ore di studio	Numero di studenti
1	10
2	30
3	30
4	20
5	10

(a) Determinare la media e la varianza campionaria;
(b) Determinare l'intervallo di confidenza al 95% per il numero medio di ore dedicate alla preparazione dell'esame di CPS da tutti gli studenti (si ricorda il valore $z_{0.95} = 1.96$).

28/06/2011

(1) Dieci centralini ricevono ciascuno un numero di telefonate al minuto avente densità di Poisson di media 2.

(a) Qual è la probabilità che il primo di questi centralini riceva almeno 3 telefonate in un minuto?
La variabile $X_1 =$"numero di telefonate al primo centralino" ha densità di Poisson di parametro 2. Si ha quindi

$$\begin{aligned} P(X_1 \geq 3) &= 1 - P(X_1 = 0) - P(X_1 = 1) - P(X_1 = 2) \\ &= 1 - e^{-2}\Big(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!}\Big) = 1 - 5e^{-2} \\ &= 0.3233. \end{aligned}$$

(b) Qual è la probabilità che almeno tre di questi centralini ricevano 3 o più telefonate ciascuno in un minuto?
La variabile $Z =$"numero di centralini che ricevono almeno 3 telefonate" è una variabile binomiale $B(10, p)$, con $p = 0.3233$. Si ha quindi

$$\begin{aligned} P(Z \geq 3) &= 1 - P(Z = 0) - P(Z = 1) - P(Z = 2) \\ &= 1 - (1-p)^{10} - 10p(1-p)^9 - \binom{10}{2} p^2 (1-p)^8 \\ &= 0.677. \end{aligned}$$

(c) Sapendo che esattamente 5 centralini hanno ricevuto 3 o più telefonate, qual è la probabilità che esattamente 2 dei primi 4 abbiano ricevuto 3 o più telefonate?
Possiamo risolvere questo quesito con l'ausilio di una variabile ipergeometrica. Pensiamo ai 10 centralini come 10 palline, di cui 5 rosse (corrispondenti ai centralini che hanno ricevuto almeno 3 telefonate) e 5 bianche (i rimanenti). Peschiamo 4 palline senza rimpiazzo (i primi 4 centralini) e ci chiediamo qual è la probabilità che esattamente 2 siano bianche. La risposta è

$$\frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}} = \frac{10}{21} = 0.476.$$

Si poteva comunque rispondere a questa domanda anche utilizzando la probabilità condizionale (e la formula di Bayes).

(2) Siano X ed Y due variabili aleatorie definite sullo stesso spazio di probabilità la cui densità congiunta è espressa dalla seguente tabella

$X \backslash Y$	-1	0	2
3	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{8}$
4	$\frac{1}{24}$	$\frac{1}{12}$	$\frac{1}{8}$
6	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{4}$

(a) Determinare $P(X + Y \geq 6)$.
Si vede facilmente che $X + Y \geq 6$ nei 3 casi (4,0), (4,2) e (6,2). Si ha quindi

$$P(X + Y \geq 6) = \frac{1}{8} + \frac{1}{6} + \frac{1}{4} = \frac{13}{24}.$$

(b) Determinare le densità marginali di X ed Y;
Basta fare la somma delle righe e delle colonne della tabella. Si ha

$$p_X(k) = \begin{cases} \frac{1}{4} & \textit{se } k = 3, 4; \\ \frac{1}{2} & \textit{se } k = 6; \\ 0 & \textit{altrimenti} \end{cases}$$

e

$$p_Y(k) = \begin{cases} \frac{1}{6} & \textit{se } k = -1; \\ \frac{1}{3} & \textit{se } k = 0; \\ \frac{1}{2} & \textit{se } k = 2; \\ 0 & \textit{altrimenti} \end{cases}$$

(c) Stabilire se X ed Y sono indipendenti;
Si vede facilmente che la densità congiunta $p(x, y)$ è uguale al prodotto delle marginali $p_X(x) \cdot p_Y(y)$ PER OGNI coppia di valori x ed y e quindi le variabili X ed Y sono indipendenti.

(d) Determinare la media di $X + Y$.
Calcoliamo la media di X e di Y utilizzando le densità marginali. Abbiamo

$$E[X] = 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} + 6 \cdot \frac{1}{2} = \frac{19}{4},$$

$$E[Y] = -1\frac{1}{6} + 0 \cdot \frac{1}{3} + 2 \cdot \frac{1}{2} = \frac{5}{6}$$

e quindi

$$E[X + Y] = E[X] + E[Y] = \frac{67}{12} = 5.58.$$

(e) Determinare la varianza di $X - Y$.
Siccome X e $-Y$ sono indipendenti abbiamo

$$\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(-Y) = \mathrm{Var}(X) + (-1)^2\mathrm{Var}(Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Calcoliamo le medie dei quadrati per poter determinare le varianza di X ed Y. Abbiamo

$$E[X^2] = (16 + 9)\frac{1}{4} + 36\frac{1}{2} = \frac{97}{4}$$

e

$$E[Y^2] = \frac{1}{6} + 4\frac{1}{2} = \frac{13}{6}.$$

Di conseguenza $\mathrm{Var}(X) = E[X^2] - E[X]^2 = \frac{97}{4} - \frac{361}{16} = \frac{27}{16}$ *e* $\mathrm{Var}(Y) = E[Y^2] - E[Y]^2 = \frac{13}{6} - \frac{25}{36} = \frac{53}{36}$. *Concludiamo che*

$$\mathrm{Var}(X - Y) = \frac{27}{16} + \frac{53}{36} = 3.16.$$

(3) Sia X una variabile continua uniforme nell'intervallo $[0, 1]$.

(a) Determinare $E[X^2]$ e $E[X^4]$.
Ricordando come si calcola la media di una variabile trasformata $\phi(X)$ abbiamo

$$E[X^2] = \int_0^1 t^2\, dt = \Big[\frac{t^3}{3}\Big]_0^1 = \frac{1}{3}$$

e similmente

$$E[X^4] = \int_0^1 t^4\, dt = \Big[\frac{t^5}{5}\Big]_0^1 = \frac{1}{5}.$$

In alternativa si poteva determinare la densità di X^2 e di X^4, ma sarebbe stato più complesso. Ad esempio ricordando quanto fatto in aula per calcolare la densità di una variabile trasformata abbiamo che la densità di X^2 è $f_{X^2}(t) = \frac{1}{2}t^{-\frac{1}{2}}$ (nell'intervallo $[0,1]$ e quindi

$$E[X^2] = \int_0^1 t\frac{1}{2}t^{-\frac{1}{2}}\,dt = \frac{1}{2}\Big[\frac{t^{3/2}}{3/2}\Big]_0^1 = \frac{1}{2}\frac{2}{3} = \frac{1}{3}.$$

(b) Siano $X_1, \dots, X_{100}$ variabili continue uniformi nell'intervallo $[0,1]$ indipendenti. Determinare

$$P(X_1^2 + \cdots + X_{100}^2 > 25)$$

Determiniamo la varianza delle variabili X_i^2. Abbiamo

$$\mathrm{Var}(X_i^2) = E[X_i^4] - E[X_i^2]^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$

La variabile $S_{100} = X_1^2 + \cdots + X_{100}^2$, per il teorema limite centrale, è approssimata da una variabile normale di media $100 \cdot \frac{1}{3} = 33.33$ e varianza $100 \cdot \frac{4}{45} = 8.88$. Abbiamo quindi

$$P(X_1^2 + \cdots + X_{100}^2 > 25) = P(\zeta_0 > \frac{25 - 33.33}{\sqrt{8.88}}) = P(\zeta_0 > -2.79) = 0.997.$$

19/07/2011

(1) Sia X la variabile aleatoria data dal numero di lanci di una moneta necessari per ottenere testa 2 volte. Si denoti con p la densità di X.
  (a) Determinare $p(1)$, $p(2)$ e $p(3)$;
  (b) determinare $p(k)$ per ogni $k > 0$;
  (c) verificare che X è somma di due variabili indipendenti aventi la stessa densità geometrica modificata;
  (d) calcolare $E[X]$;

Soluzione.
  (a) Evidentemente $p(1) = 0$ in quanto non possiamo ottenere 2 teste con 1 lancio. $p(2) = 1/4$ in quanto la probabilità di ottenere 2 teste su 2 lanci è $1/4$. Riguardo $p(3)$, ci sono 2 modi in cui otteniamo testa per la seconda volta al terzo lancio (TCT e CTT) sugli 8 possibili e quindi $p(3) = 2/8 = 1/4$.
  (b) Sia Y la variabile "numero di teste nei primi $k-1$ lanci" e Z la variabile $B(1, 1/2)$ che fa 1 se il k-esimo lancio dà testa. Si ha
$$\{X = k\} = \{Y = 1\} \cap \{Z = 1\}.$$
Si ha chiaramente che Y è binomiale $B(k-1, 1/2)$ e che Y e Z sono indipendenti. Abbiamo quindi
$$p(k) = P(X = k) = P(Y = 1) \cdot P(Z = 1) = (k-1)(1/2)^{k-1} \cdot (1/2) = (k-1)(1/2)^k.$$
  (c) Se denotiamo con X_1 il numero di lanci necessari per ottenere la prima testa e e con X_2 il numero di lanci necessari per ottenere la seconda testa dopo aver ottenuto la prima, abbiamo chiaramente $X = X_1 + X_2$ e che le due variabili X_1 e X_2 sono geometriche modificate di parametro $1/2$.
  (d) Si ha $E[X] = E[X_1] + E[X_2] = \frac{1}{1/2} + \frac{1}{1/2} = 4$.

(2) Un impianto d'allarme dà falsi allarmi in intervalli di tempo che seguono una legge esponenziale di media 1 anno.
  (a) Calcolare la probabilità che nel prossimo anno ci sia qualche falso allarme;
  (b) sapendo che in un anno ci sono stati 2 falsi allarmi, qual è la probabilità che nell'anno successivo ci sia almeno un falso allarme?
  (c) Due case montano questo impianto d'allarme. Qual è la probabilità che almeno 1 dia un falso allarme nei prossimi 6 mesi? E tutti e 2?

Soluzione.
  (a) Sia T la variabile "tempo necessario (in anni) per avere un falso allarme". (Siccome la variabile T soddisfa la proprietà di mancanza di memoria non è necessario specificare l'istante iniziale). Per ipotesi la variabile T è esponenziale di parametro $\lambda = 1$ e quindi la sua funzione di ripartizione è data da $F(t) = 1 - e^{-t}$. Sia A l'evento $A = \{$ Si verifica almeno un falso allarme nel prossimo anno $\}$. Si ha chiaramente $A = \{T \leq 1\}$ e quindi
$$P(A) = P(T \leq 1) = F(1) = 1 - e^{-1} = 0.632.$$
  (b) Per la mancanza di memoria della variabile esponenziale questa probabilità è uguale a quella calcolata nel punto (a).
  (c) Siano T_1 e T_2 le variabili "tempo necessario (in anni) per avere un falso allarme" delle 2 abitazioni. Possiamo a questo punto ricordare che $\min(T_1, T_2)$ è ancora esponenziale di parametro 1+1=2 e quindi
$$P(\min(T_1, T_2) < \frac{1}{2}) = 1 - e^{-2\frac{1}{2}} = 0.632.$$

Riguardo il massimo possiamo ricordare la formula da lezione oppure ricavarcela osservando che

$$P(\max(T_1,T_2)<\frac{1}{2})=P(T_1<\frac{1}{2},T_2<\frac{1}{2})=P(T_1<\frac{1}{2})P(T_2<\frac{1}{2})=(1-e^{-\frac{1}{2}})^2=0.159.$$

(3) Si vuole fare un test per capire se un nuovo medicinale sia efficace in almeno il 70% dei casi. Provando tale medicinale su 78 pazienti determinare in quali casi si accetta l'efficacia del medicinale con significatività al 5% e all'1%.

Soluzione.

Siamo nella tipica situazione in cui è più opportuno effettuare un test ad una coda. Formuliamo quindi l'ipotesi $H_0 : \mu \geq 0.70$, dove μ è la probabilità che la medicina sia efficace su un paziente scelto a caso. Effettuando un test su 78 pazienti, nell'ipotesi $\mu = 0.70$ abbiamo che la variabile $\bar{X}_{78}$ avrà media 0.70 e scarto quadratico medio $\sqrt{\frac{0.70\cdot 0.30}{78}} = 0.052$. Abbiamo quindi che $\bar{X}_{78} \sim N(0.7,(0.052)^2)$ e quindi con probabilità 95% avremo

$$\bar{X}_{78} > 0.70 - 1.64 \cdot 0.052 = 0,615$$

e quindi accettiamo l'ipotesi con significatività al 5% se il medicinale è efficace su $0,615 \cdot 78 = 48$ pazienti.

Con significatività all'1% accetteremo l'ipotesi se il numero di pazienti che hanno avuto giovamento è almeno $78(0.70 - 2.33 \cdot 0.052) = 46$.

05/09/2011

(1) In un quiz vengono poste ad un concorrente 5 domande aventi 3 possibili risposte ciascuna. Ogni volta che commette un errore deve ricominciare daccapo rispondendo alla prima domanda, finché non le indovina tutte. Si supponga che il concorrente risponda sempre a caso tenendo a mente però le risposte già date.
 (a) Qual è il numero massimo e il numero minimo di domande che possono essere fatte?
 (b) Detta X_i la variabile "numero di errori alla domanda i", determinare la densità delle variabili X_i;
 (c) esprimere la variabile $Y =$ "numero totale di domande effettuate" in funzione delle X_i;
 (d) determinare la media e la varianza della variabile Y;
 (a) Il numero minimo di domande è 5 (caso in cui vengono indovinate tutte subito). Il numero massimo si ottiene considerando il caso in cui ogni domanda viene sbagliata 2 volte prima di essere indovinata. In questo caso ognuno dei 2 errori alla domanda i viene preceduto necessariamente da $i-1$ domande indovinate e quindi in tutto abbiamo 1+1+2+2+3+3+4+4+5+5+5=35 domande, dove l'ultimo 5 rappresenta l'ultima sequenze di 5 domande indovinate.
 (b) Le variabili X_i possono assumere i valori 0,1,2. "$X_i = 0$" corrisponde al caso in cui si risponde correttamente alla domanda i al primo tentativo, e quindi $P(X_i = 0) = 1/3$. "$X_i = 1$" corrisponde al caso in cui si risponde correttamente alla domanda i al secondo tentativo. In questo caso si ha $P(X_i = 1) = P(\text{sbaglio al primo tentativo e indovino al secondo}) = P(\text{sbaglio al primo})P(\text{ indovino al secondo} \mid \text{sbaglio al primo}) = \frac{2}{3}\frac{1}{2} = 1/3$. Per complementarità abbiamo anche $P(X_i = 2) = 1/3$ e quindi concludiamo che le variabili X_i sono tutte variabili uniformi sull'insieme $\{0, 1, 2\}$. (Che legame c'è secondo te con una variabile successo-insuccesso senza rimpiazzo?)
 (c) Come già osservato al punto (a) un errore alla domanda i comporta la formulazione di $i - 1$ domande indovinate che precedono l'errore stesso, a cui bisogna aggiungere le ultime 5 domande che vengono senz'altro indovinate. Abbiamo quindi

$$Y = 5 + \sum_{i=1}^{5} (iX_i).$$

 (d) Si ha

$$E[Y] = 5 + \sum iE[X_i] = 5 + \sum i = 20.$$

Si ottiene facilmente $Var(X_i) = 2/3$ e quindi

$$\mathrm{Var}(Y) = \mathrm{Var}(5) + \sum_{i=1}^{5} \mathrm{Var}(iX_i) = 0 + \sum_{i=1}^{5} i^2 \mathrm{Var}(X_i) = \frac{110}{3} = 36,7.$$

(2) Siano

$$f(t) = \begin{cases} 0 & t \leq 1 \\ at^2 + 3t + 3a & 1 < t \leq 3 \\ 0 & t > 3 \end{cases} \quad F(t) = \begin{cases} 0 & t \leq 1 \\ bt^3 - 6bt^2 + 9bt + 1 & 1 < t \leq 3 \\ 1 & t > 3 \end{cases}$$

 (a) Determinare, se esistono, valori dei parametri a e b in modo che $f(t)$ e $F(t)$ siano rispettivamente densità e funzione di ripartizione di una variabile continua X;
 (b) determinare in tal caso $E[X]$ e $P(X > 2)$;

(a) Per essere una densità, $f(t)$ deve sempre essere ≥ 0 e deve avere integrale 1. Vediamo quest'ultima condizione

$$\begin{aligned}\int_{-\infty}^{+\infty} f(t)\,dt &= \int_1^3 (at^2+3t+3a)dt \\ &= \Big[a\frac{t^3}{3}+3\frac{t^2}{2}+3at\Big]_1^3 \\ &= \frac{44}{3}a+12.\end{aligned}$$

Dovendo essere tale integrale pari ad 1 ne deduciamo che $a=-\frac{3}{4}$. Per tale valore di a si verifica che $f(t)\geq 0$ per ogni t in quanto, ad esmpio, tra 1 e 3 abbiamo l'equazione di una parabola con concavità verso il basso che si annulla agli estremi.
A questo punto basta imporre che $F(t)$ sia la derivata di $f(t)$ ottenendo il valore $b=-\frac{1}{4}$.

(b) Si ha

$$E[X]=\int_1^3 t\Big(-\frac{3}{4}t^2+3t-\frac{9}{4}\Big)dt=\Big[-\frac{3}{16}t^4+t^3-\frac{9}{8}t^2\Big]_1^3=2,$$

mentre

$$P(X>2)=1-F(2)=1+\frac{1}{4}2^3-6\frac{1}{4}2^2+\frac{9}{4}2-1=0.5$$

(3) In un gioco vince che ottiene il punteggio più alto sommando i valori ottenuti in 100 lanci di un dado.
(a) Se il nostro avversario ha ottenuto 360 punti che probabilità abbiamo di vincere la partita?
(b) Giocando 5 partite qual è la probabilità di ottenere più di 360 punti esattamente due volte?

(a) Sia X_i = "risultato dell'i-esimo lancio". Si ha $E[X_i]=7/2$ e $E[X_i^2]=\frac{1}{6}(1+4+9+16+25+36)=91/6$ e quindi $\mathrm{Var}(X_i)=E[X_i^2]-E[X_i]^2=\frac{91}{6}-\frac{49}{4}=\frac{35}{12}$. Se poniamo quindi $S_{100}=X_1+X-2+\cdots+X_{100}$ sappiamo che S_{100} è approssimativamente normale di media $100\cdot\frac{7}{2}=354$ e varianza $100\cdot\frac{35}{12}=291,7$. Effettuando una correzione di continuità abbiamo quindi

$$P(S_{100}>360)=P(\zeta_0>\frac{360.5-350}{\sqrt{291,7}})=P(\zeta_0>0.61)=1-P(\zeta_0<0.61)=1-0.7257=0.2743$$

(b) In ogni partita abbiamo probabilità 0.27 di ottenere più di 360 punti, indipendentemente l'una dall'altra. Il numero di volte che si ottengono più di 360 punti è quindi una variabile binomiale $B(5,0.27)$ e quindi la probabilità richiesta è

$$\binom{5}{2}(0.27)^2(0.73)^3=0,283.$$

26/09/2011

(1) Dopo aver mescolato un mazzo di 40 carte Arturo scopre una carta per volta finché non compare l'asso di denari. Basilio mescola un altro mazzo di carte e e scopre una carta. Se non è l'asso di denati rimischia tutte le carte e ne riscopre una. Ripete l'operazione finché non trova l'asso di denari. Siano X ed Y le variabili "numero di carte scoperte da Arturo e Basilio" rispettivamente.

(a) Determinare le densità delle variabili X ed Y;

(b) Determinare la media di X ed Y;

(c) Determinare la funzione di ripartizione di $\max(X, Y)$.

Soluzione. (a) nel caso di Arturo abbiamo evidentemente che la probabilità che l'asso di denari sia la i-esima carta è $\frac{1}{40}$ dato che le carte sono state mischiate uniformemente. Abbiamo quindi che la densità di X è la densità uniforme su $\{1, 2, \dots, 40\}$.

$$p_X(i) = \begin{cases} \frac{1}{40} & \text{se } i = 1, 2, \dots, 40 \\ 0 & \text{altrimenti.} \end{cases}$$

La variabile Y è un tempo di primo successo in uno schema successo-insuccesso con ripetizione. Quindi Y ha densità geometrica modificata di parametro $\frac{1}{40}$ (probabilità di avere un successo al primo tentativo). Si ha quindi

$$p_Y(i) = \begin{cases} \frac{1}{40}\left(\frac{39}{40}\right)^{i-1} & \text{se } i = 1, 2, \dots, \\ 0 & \text{altrimenti.} \end{cases}$$

(b) La media di X è data da

$$E[X] = \sum_{i=1}^{40} i p_X(i) = \frac{1}{40}(1 + 2 + \cdots + 40) = 20.5;$$

La media di una variabile geometrica modificata di parametro p à data da $1/p$ e quindi in questo caso abbiamo $E[Y] = 40$.

(c) Sia $Z = \max(X, Y)$. Le funzioni di ripartizione di X ed Y sono date da (per valori di k interi positivi)

$$F_X(k) = \begin{cases} \frac{k}{40} & \text{se } i = 1, 2, \dots, 40 \\ 1 & \text{se } k > 40 \end{cases}$$

e

$$F_Y(k) = 1 - \left(\frac{39}{40}\right)^k.$$

Sfruttando il fatto che X ed Y sono chiaramente indipendenti abbiamo

$$\begin{aligned} F_Z(k) &= P(\max(X, Y) \le k) \\ &= P(X \le k), Y \le k) \\ &= F_X(k)F_Y(k) \\ &= \begin{cases} \frac{k}{40}\left(1 - \left(\frac{39}{40}\right)^k\right) & \text{se } i = 1, 2, \dots, 40 \\ 1 - \left(\frac{39}{40}\right)^k & \text{se } k > 40 \end{cases} \end{aligned}$$

(2) Un vulcano ha 2 crateri. Essi hanno un tempo di rottura, con conseguente eruzione, di densità esponenziale di media 50 e 100 anni rispettivamente, indipendenti l'uno dall'altro.

(a) Determinare la densità della variabile "tempo di eruzione del vulcano";

(b) Determinare la probabilità che ci sia (almeno) un'eruzione nei prossimi 100 anni;

(c) Determinare la probabilità che entrambi i crateri diano origine ad almeno una eruzione nei prossimi 70 anni.

Soluzione. (a) Siano X ed Y le variabili tempo di eruzione dei due vulcani. Per ipotesi quindi X ha densità esponenziale di parametro $\frac{1}{50}$ e Y ha densità esponenziale di parametro $\frac{1}{100}$. La variabile tempo di eruzione del vulcano è data da $Z = \min(X, Y)$. Ricordando (o rifacendo) l'esempio visto in classe (vedi anche il libro a pagina 96) concludiamo che Z ha densità esponenziale di parametro $1/50 + 1/100 = 3/100$.

(b) Bisogna calcolare $P(Z \leq 100) = F_Z(100) = 1 - e^{-\frac{3}{100}100} = 1 - e^{-3} = 0.95$. (c) In questo caso bisogna considerare la variabile $\max(X, Y)$ e calcolare

$$P(\max(X, Y) \leq 70).$$

Si ha $P(\max(X, Y) \leq 70) = P(X \leq 70)P(Y \leq 70) = (1 - e^{-\frac{1}{50}70})(1 - e^{-\frac{1}{100}70}) = 0.379$

(3) Siano $X_1, \ldots, X_{100}$ variabili indipendenti uniformi su $\{1, 2, 3, 4, 5, 6\}$ e $Y_1, \ldots, Y_{100}$ variabili indipendenti uniformi sull'intervallo $[1, 6]$. Poniamo $X = X_1 + \cdots + X_{100}$ e $Y = Y_1 + \cdots Y_{100}$.

(a) Determinare la media e la varianza di X e di Y

(b) Determinare $P(Y > 370)$;

(c) Determinare $P(X > 370)$.

Soluzione. La media delle X_i è $E[X_i] = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$. La media delle Y_i è il punto medio dell'intervallo $[1, 6]$ e quindi è sempre $\frac{7}{2}$. Per calcolare la varianza delle X_i determiniamo la media di X_i^2 abbiamo: $E[X_i^2] = \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = 91/6$ e quindi

$$\mathrm{Var}(X_i) = E[X_i^2] - E[X_i]^2 = \frac{91}{6} - \frac{49}{4} = \frac{182 - 147}{12} = \frac{35}{12}.$$

La varianza di una variabile uniforme nell'intervallo $[a, b]$ è data da $\frac{(b-a)^2}{12}$ e quindi nel nostro caso abbiamo $\mathrm{Var}(Y_i) = \frac{25}{12}$. Abbiamo quindi $E[X] = 100 \cdot E[X_1] = 350$, $E[Y] = 100 \cdot E[Y_1] = 350$, $\mathrm{Var}(X) = 100 \cdot \mathrm{Var}(X_1) = \frac{875}{3}$, $\mathrm{Var}(Y) = 100 \cdot \mathrm{Var}(Y_1) = \frac{625}{3}$.

(b)Per il teorema limite centrale abbiamo $Y \sim N(350, \frac{625}{3})$ e quindi

$$P(Y > 370) = P(\zeta_0 > \frac{370 - 350}{\sqrt{625/3}}) = P(\zeta_0 > 1.38) = 1 - \Phi(1.38) = 1 - 0.916 = 0.084.$$

(c) Similmente abbiamo $X \sim N(350, \frac{875}{3})$ e facendo la correzione di continuità visto che X è una variabile discreta abbiamo

$$P(X > 370) = P(\zeta_0 > \frac{370.5 - 350}{\sqrt{875/3}}) = P(\zeta_0 > 1.20) = 1 - \Phi(1.20) = 1 - 0.885 = 0.115.$$

17/01/2012

(1) Abbiamo un dado e un sacchetto contenente 4 palline blu e 2 rosse. Lanciamo il dado e denotiamo con X il risultato. Poi peschiamo X palline dal sacchetto e denotiamo con Y il numero di palline rosse estratte.
  (a) Determinare la densità di X;
  (b) Calcolare $P(Y = 1|X = 3)$;
  (c) Determinare la densità di Y;
  (d) Stabilire se X ed Y sono indipendenti;
  (e) Sapendo di aver pescato entrambe le palline rosse determinare la probabilità di aver ottenuto 5 nel lancio del dado.

(2) Si consideri la seguente funzione

$$f(t) = \begin{cases} t & \text{se } 0 \leq t < 1 \\ 2 - t & \text{se } 1 \leq t < 2 \\ 0 & \text{altrimenti} \end{cases}$$

  (a) Verificare che $f(t)$ è la densità di una variabile aleatoria continua X.
  (b) Deteminare la funzione di ripartizione di X.
  (c) Sapendo che X è la variabile somma di 2 variabili indipendenti uniformi nell'intervallo $[0, 1]$, determinare qual è la probabilità che la somma di due numeri casuali compresi tra 0 e 1 sia maggiore di 1.5.

(3) In un chiosco per piadine un sacco di farina dura mediamente 3 giorni con uno scarto quadratico medio di 6 ore.
  (a) Quanto dura mediamente una fornitura di 100 sacchi di farina? Con quale varianza?
  (b) Con quale probabilità una fornitura di 120 sacchi di farina può bastare per un anno di attività?

20/02/2012

(1) Lanciamo un dado e una moneta simultaneamente per 3 volte. Denotiamo con X la variabile aleatoria che conta il numero di teste uscite e Y la variabile aleatoria che conta il numero di volte che il dado ha dato come risultato 1 o 3.
   (a) Determinare le densità delle variabili X ed Y;
   (b) Determinare la funzione di ripartizione di $Z = \max(X, Y)$.
   (c) Determinare la media e la varianza di $X + Y$.
   (d) Determinare $P(X + Y = 5)$.

(2) Si considerino le seguenti variabili aleatorie indipendenti: X, di densità esponenziale di media 1, ed Y, di densità uniforme nell'intervallo $[0, 2]$.
   (a) Determinare $P(X > 1)$ e $P(Y > 1)$;
   (b) Determinare la probabilità che la variabile $\min(X, Y)$ sia maggiore di 1;
   (c) Determinare la probabilità che la variabile $\max(X, Y)$ sia maggiore di 2.

(3) Si vuole stimare la media dei voti verbalizzati all'esame di CPS. Si sa che su un campione di 50 studenti il voto medio è stato 24 con uno scarto quadratico medio di 2.
   (a) Determinare l'intervallo di confidenza al 95% (si ricorca $z_{0.95} = 1.96$;
   (b) Supponendo sempre una media di 24 e uno scarto quadratico medio di 2, quanti studenti deve contenere il campione per ottenere come intervallo di confidenza al 95% l'intervallo $[23.9, 24.1]$?

Soluzione (1) (a) X ha densità binomiale $B(3, \frac{1}{2})$ mentre Y ha densità binomiale $B(3, \frac{1}{3})$. Abbiamo quindi:

$$p_X(k) = \begin{cases} \frac{1}{8} & \text{se } k = 0, 3 \\ \frac{3}{8} & \text{se } k = 1, 2 \\ 0 & \text{altrimenti} \end{cases}$$

e

$$p_Y(h) = \begin{cases} \frac{8}{27} & \text{se } k = 0 \\ \frac{12}{27} & \text{se } k = 1 \\ \frac{6}{27} & \text{se } k = 2 \\ \frac{1}{27} & \text{se } k = 3 \\ 0 & \text{altrimenti} \end{cases}$$

(b) $P(Z \le 0) = P(X \le 0)P(Y \le 0) = \frac{1}{27}$;
$P(Z \le 1) = P(X \le 1)P(Y \le 1) = \frac{1}{2} \cdot \frac{20}{27} = \frac{10}{27}$;
$P(Z \le 2) = P(X \le 2)P(Y \le 2) = \frac{7}{8} \cdot \frac{26}{27} = \frac{91}{108}$;
$P(Z \le 3) = P(X \le 3)P(Y \le 3) = 1$.

(c)$E[X + Y] = E[X] + E[Y] = \frac{3}{2} + 1 = 2.5$;
$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = \frac{3}{4} + \frac{2}{3} = \frac{17}{12}$.

(d) $P(X + Y = 5) = P(X = 2)P(Y = 3) + P(X = 3)P(Y = 2) = \frac{1}{24}$.

(2) (a) La funzione di ripartizione di X è $1 - e^{-t}$ per $t > 0$ e quindi

$$P(X > 1) = e^{-1} = 0.37$$

Mentre $P(X > 1)$ è chiaramente 0.5

(b) E' semplicemente il prodotto $0.5 \cdot 0.37 = 0,19$

(c) E' data dalla probabilitÃ che X sia maggiore di 2 dato che Y non lo puó esserlo e quindi abbiamo $e^{-2} = 0.13$

(3) (a)L'intervallo di confidenza ha centro 24 e semiampiezza $1.96 * 2/\sqrt{50} = 0.55$. Possiamo quindi concludere che l'intercallo di confidenza al 95% per la media dei voti è $[23.45, 24.55]$.

(b) Per ottenere una semiampiezza pari a 0.1 dobbiamo imporre

$$1.96 \cdot 2/\sqrt{n} = 0.1$$

da cui ricaviamo che n deve essere almeno 1536. Per avere un'ottima precisione nella stima abbiamo quindi bisogno di un campione molto numeroso.

01/06/2012

(1) Lanciamo una moneta ripetutamente e poniamo $X_i = 1$ se l'i-esimo lancio ha dato testa e $X_i = 0$ altrimenti. Consideriamo le variabili X e Y date da $X = \min\{i : X_{i-1} = X_i\}$ e $Y = \min\{i : X_{i-1} = X_i = 1\}$. Ad esempio ottenendo una sequenza 1001011 abbiamo $X = 3$ e $Y = 7$.
  (a) Determinare $P(X_i = X_{i-1}$ per ogni $i > 1$.
  (b) Determinare le densità della variabile X;
  (c) Determinare la media della variabile X;
  (d) (*) Detto a_i il numero di sequenze binarie di lunghezza i in cui 1 non compare mai 2 volte consecutivamente, dimostrare che $a_{i+2} = a_{i+1} + a_i$, per $i \geq 1$.
  (e) Determinare esplicitamente a_1, a_2, a_3, a_4, a_5.
  (f) Dimostrare che la densità di Y è data da $p_Y(k) = \frac{a_{k-3}}{2^k}$ per $k > 3$
  (g) Utilizzare (e) ed (f) per determinare $p_Y(k)$ per $k \leq 8$;
  (h) (*) Dimostrare che mediamente occorrono almeno 5 lanci per ottenere 2 teste consecutive.

Soluzione (1) (a)Ad ogni lancio, con eccezione del primo, la probabilità di ottenere un risultato uguale al precedente è sempre $\frac{1}{2}$ (b) Siccome la probabilità che un lancio sia uguale al precedente è indipendente dai lanci precedenti, segue da (a) che $X - 1 \sim \tilde{G}(\frac{1}{2})$.
(c) Essendo $X - 1 \sim \tilde{G}(\frac{1}{2})$ abbiamo che $E[X] = E[X - 1] + 1 = 2 + 1 = 3$.
(d) Consideriamo le sequenze di lunghezza $i + 2$ in cui non compaiono mai 2 teste consecutive. Le suddividiamo in 2 sottoinsiemi, quelle che finiscono con testa e quelle che finiscono con croce. Quelle che finiscono con croce hanno nei primi $i+1$ posti una qualunque sequenza di teste e croci senza teste consecutive e quindi sono a_{i+1}. Quelle che finiscono con testa hanno necessariamente croce nel posto $i + 1$ e una qualunque sequenze nel primi i posti e quindi sono proprio a_i.
(e) Si ha facilmente $a_1 = 2$ e $a_2 = 3$. Dal punto precedente otteniamo quindi $a_3 = 5$, $a_4 = 8$, $a_5 = 13$.
(f) Il numero di sequenze possibili di lunghezza k è 2^k. Le sequenze di lunghezza k ($k \geq 3$) che hanno 2 teste consecutive per la prima volta nei posti $k-1$ e k sono sequenze che finiscono con CTT e hanno nei primi $k - 3$ posti una qualunque sequenza che non ha mai 2 teste consecutive. Sono quindi a_{k-3}. Otteniamo quindi $p(k) = \frac{a_{k-3}}{2^k}$ per $k > 3$ (g)

$$p_Y(2) = \frac{1}{4},\ p_Y(3) = \frac{1}{8}\ p_Y(4) = \frac{a_1}{2^4} = \frac{1}{8},\ p_Y(5) = \frac{a_2}{2^5} = \frac{3}{32},\ p_Y(6) = \frac{5}{64},\ p_Y(7) = \frac{1}{16},\ p_Y(8) = \frac{13}{256}.$$

(h) Dal punto (g) abbiamo che $P(Y \geq 9) = 1 - (p(2) + \cdots + p(8)) = 0.215$ e quindi

$$\begin{aligned} E[Y] &= \sum_{k=2}^{\infty} kp(k) = \sum_{k=2}^{8} kp(k) + \sum_{k=9}^{\infty} kp(k) \geq \sum_{k=2}^{8} kp(k) + \sum_{k=9}^{\infty} 9p(k) \\ &= 2\frac{1}{4} + 3\frac{1}{8} + 4\frac{1}{8} + \cdots + 8\frac{13}{256} + 9 \cdot 0.215 = 5.09. \end{aligned}$$

(2) Una squadra di calcio vince con probabilità $\frac{1}{2}$ e pareggia con probabilità $\frac{1}{4}$.
  (a) Giocando 30 partite in un campionato, qual è la probabilità che ne vinca esattamente 15?
  (b) E che ne vinca più di 15?
  (c) Qual è la probabilità che faccia almeno 50 punti?

(2) (a) Il numero X di vittorie segue una densità binomiale $B(30, \frac{1}{2})$ e quindi

$$P(X = 15) = \binom{30}{15}(1/2)^{30} = 0,144.$$

(b) Approssimiamo la variabile X con una variabile normale $\zeta \sim N(15, \frac{15}{2})$ e calcoliamo quindi

$$P(X > 15) = P(\zeta \geq 15.5) = P(\zeta_0 > \frac{15.5 - 15}{\sqrt{7.5}} = P(\zeta_0 > 0.18) = 1 - \Phi(0.18) = 1 - 0.57142 = 0.42858.$$

In alternativa, senza ricorrere all'approssimazione normale si poteva risolvere sfruttando la simmetria del problema: la probabilità di ottenere più di 15 vittorie è uguale alla probabilità di ottenere meno di 15 vittorie. E siccome $P(X < 15) + P(X = 15) + P(X > 15) = 1$ otteniamo

$$P(X > 15) = \frac{1 - P(X = 15)}{2} = \frac{1 - 0.144}{2} = 0.428,$$

esattamente lo stesso risultato!

(c) Detto Y_i i punti conquistati in ogni partita abbiamo che la densità delle Y_i è la seguente

$$p_Y(k) = \begin{cases} \frac{1}{2} & k = 3 \\ \frac{1}{4} & k = 1 \\ \frac{1}{4} & k = 0 \\ 0 & \textit{altrimenti} \end{cases}$$

Abbiamo quindi $E[Y_i] = 3\frac{1}{2} + 1\frac{1}{4} = \frac{7}{4}$ e $E[Y_i^2] = 9\frac{1}{2} + 1\frac{1}{4} = \frac{19}{4}$ e quindi $\mathrm{Var}(Y_i) = E[Y_i^2] - E[Y_i]^2 = \frac{19}{4} - \frac{49}{16} = \frac{27}{16}$. La variabile somma di punti $S_{30} = Y_1 + \cdots + Y_{30}$ per il teorema del llimite centrale è unva variabile approssimativamente normale di media $E[S_{30}] = 30\frac{7}{4} = 52.5$ e varianza $\mathrm{Var}(S_{30}) = 30 \cdot \frac{27}{16} = 50.625$. Utilizzando ancora la correzione di continuità abbiamo

$$P(S_{30} \geq 50) = P(\zeta > 49.5) = P(\zeta_0 > \frac{49.5 - 52.5}{\sqrt{50,625}} = P(\zeta_0 > -0.42) = \Phi(0.42) = 0.66276$$

03/07/2012

(1) Si lancia un dado per 3 volte e poniamo, per $i = 1, 2, 3, 4, 5, 6$

$$X_i = \text{numero di lanci in cui si è ottenuto un risultato } \leq i.$$

(a) Si determinino le densità delle X_i.
(b) Stabilire per quali coppie (i, j), con $i < j$ le variabili X_i e X_j sono indipendenti.
(c) Verificare che $P(X_3 = 1, X_5 = 3) = \frac{1}{6}$;
(d) (*) Determinare $\mathrm{Cov}(X_3, X_5)$.
(e) Determinare $P(X_5 = 3|X_3 = 1)$.
(a) La probabilità di ottenere un risultato $\leq i$ è $i/6$ od ogni lancio, indipendentemnete l'uno dall'altro. Si ha quindi
$$X_i \sim B(3, i/6).$$
(b) Le variabili X_i ed X_j non sono mai indipendenti se $j < 6$. Infatti si ha $P(X_i = 1, X_j = 0) = 0$ (se $i < j$) ma chiaramente $P(X_i = 1)$ e $P(X_j = 0)$ sono entrambe non nulle. La variabile X_6, essendo costante uguale a 6, è indipendente con tutte le altre.
(c) Dobbiamo calcolare la probabilità che i 3 lanci diano sempre un risultato ≤ 5 di cui una sola volta ≤ 3. Contiamo i risultati favorevoli: risultato ≤ 3 può capitare in ognuno dei 3 lanci e per ognuno di essi ho 3 possibilità (1,2 o 3), dando quindi un fattore 9. Negli altri 2 lanci devo ottenere 4 o 5 ottenendo quindi un fattore 4. Complessivamente:
$$P(X_3 = 1, X_5 = 3) = \frac{9 \cdot 4}{6^3} = \frac{1}{6}.$$
(d) Detta $p(x, y)$ la densità congiunta delle variabili X_3 ed X_5 si ha, ragionando similmente al punto precedente: $p(3, 3) = \frac{1}{8}$
$p(3, 2) = \frac{1}{4}$
$p(3, 1) = \frac{1}{6}$
$p(3, 0) = \frac{1}{27}$
$p(2, 2) = \frac{1}{8}$
$p(2, 1) = \frac{1}{6}$
$p(2, 0) = \frac{1}{18}$
$p(1, 1) = \frac{1}{24}$
$p(1, 0) = \frac{1}{36}$
$p(0, 0) = \frac{1}{216}$
In realtà non è necessario calcolare tutta la densità congiunta: per calcolare la media di $E[X_3 \cdot X_5]$ era sufficiente calcolare solo i 5 casi in cui entrambe la variabili sono non nulle.
$$E[X_3 \cdot X_5] = \frac{1}{8} \cdot 9 + \frac{1}{4} \cdot 6 + \frac{1}{8} \cdot 4 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 2 + \frac{1}{24} \cdot 1 = 4.$$
La covarianza é quindi
$$Cov(X_3, X_5) = E[X_3 \cdot X_5] - E[X_3]E[X_5] = 4 - \frac{3}{2} \cdot \frac{5}{2} = \frac{1}{4}.$$
Era da aspettarsi una covarianza positiva?
(e) Per la definizione di probabilità condizionale abbiamo $P(X_5 = 3|X_3 = 1) = \frac{P(X_5=3, X_3=1)}{P(X_3=1} = \frac{1/6}{3/8} = \frac{4}{9}$.

(2) Si considerino le funzioni

$$f_1(t) = \begin{cases} 1 - t^2 & \text{se } t \in [-1, 1] \\ 0 & \text{altrimenti} \end{cases} \quad f_2(t) = \begin{cases} \frac{1}{2} - t & \text{se } t \in [-1, 1] \\ 0 & \text{altrimenti} \end{cases} \quad f_3(t) = \begin{cases} \frac{3}{4}(1 - t^2) & \text{se } t \in [-1, 1] \\ 0 & \text{altrimenti.} \end{cases}$$

(a) Stabilire quali tra le f_1, f_2, f_3 rappresentano la densità di una variabile aleatoria continua.
(b) Determinare media e varianza di una variabile aleatoria continua la cui densità è una delle f_i.

(3) In un'inchiesta viene chiesto a 100 studenti scelti a caso in Europa il numero di viaggi effettuato all'estero negli ultimi 5 anni. I risultati sono rappresentati nella seguente tabella

N. di Viaggi	Frequenza
0	15
1	25
2	21
3	20
4	14
≥ 5	5

(a) Determinare la media e la varianza del campione.
(b) Determinare l'intervallo di confidenza al 90% per il numero medio di viaggi effettuato da uno studente europeo.
(c) Determinare l'intervallo di confidenza al 95% per la percentuale di studenti che hanno effettuato almeno 3 viaggi.

17/09/2012

(1) Un'urna contiene n palline numerate da 1 ad n. Estraiamo tutte le palline, una per volta, senza rimpiazzo. Per $i = 1, \dots, n$ consideriamo la variabile

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-esima estratta è la pallina numero } i \\ 0 & \text{altrimenti} \end{cases}$$

(a) Per ogni i, si determini la densità della variabile X_i.
(b) Stabilire per quali coppie (i, j), con $i < j$ le variabili X_i e X_j sono indipendenti.
(c) Determinare la media di $X = X_1 + \cdots + X_n$;
(d) Per $n = 3$ determinare $\text{Var}(X)$.

(2) Sia X una variabile esponenziale di parametro $\lambda = 2$.
(a) Determinare $E[1 - 2X]$;
(b) Deteminare $\text{Var}(1 - 2X)$;
(c) Determinare la densità di $\sqrt{X}$;

(3) Sia X il numero di teste ottenute in n lanci di una moneta.
(a) Si determinino la media e la varianza di X in funzione di n.
(b) Per $n = 12$ determinare $P(X \leq 10)$.
(c) Determinare il massimo valore di n per cui $P(X \leq 10) > 0.92$.

08/01/2013

(1) Voglio scaricare dalla rete 50 file di musica, uno di seguito all'altro. Posso supporre che il tempo, espresso in secondi, che il mio computer impiega per scaricare ciascun file sia una variabile aleatoria T_k definita da $T_k := 10 + X_k$, $k = 1, \ldots, 50$, dove le X_k sono variabili aleatorie indipendenti aventi la stessa densità di media pari a 1 minuto e varianza pari a 120 sec^2. Il tempo totale impiegato per scaricare i 50 file (espresso in secondi) è quindi $T := T_1 + \cdots + T_{50}$.
   (a) Calcolare il tempo totale medio $E(T)$ e la varianza $Var(T)$.
   (b) Calcolare la probabilità che il tempo totale per scaricare i 50 file sia superiore a 1 ora.
   (c) Supponiamo che ognuno dei 50 file scaricati abbia una probabilità $p = 0.01$ di essere difettoso, indipendentemente l'uno dall'altro. Calcolare la probabilità che il numero di file difettosi sia inferiore o uguale a 2.

(2) Una trasmissione televisiva è stata vista dal 25.5% delle femmine di età superiore a 14 anni, dal 62% dei maschi di età superiore a 14 anni e dal 6% degli individui di età minore o uguale a 14 anni. Inoltre, sappiamo che la popolazione italiana è formata per il 14% da individui di età minore o uguale a 14 anni, per il 44.7% da donne di età superiore a 14 anni e per il rimanente 41.3% da maschi di età superiore a 14 anni.
   (a) Calcolare la probabilità che un individuo scelto a caso tra la popolazione italiana abbia visto la partita.
   (b) Sapendo che un individuo scelto a caso ha visto la partita, qual è la probabilità che sia femmina e di età superiore a 14 anni?

(3) Sia X una variabile aleatoria continua uniforme nell'intervallo $[0, 1]$.
   (a) Calcolare $P(-log(X) \geq 1)$.
   (b) Determinare la funzione di ripartizione di $-log(X)$ e riconoscerla.

29/01/2013

(1) Si consideri un codice a 3 cifre decimali $a_1a_2a_3$ scelte a caso e indipendentemente l'una dall'altra. Per $i = 1, 2$ poniamo

$$X_i = \begin{cases} 1 & \text{se } a_i = a_{i+1} \\ 0 & \text{altrimenti.} \end{cases}$$

(a) Stabilire la densità delle X_i.
(b) Stabilire se le variabili X_1 e X_2 sono indipendenti.
(c) Stabilire la densità di $X_1 + X_2$ e di $X_1 \cdot X_2$.

(a) le X_i sono variabili di Bernoulli, in quanto assumono solo i valori 0 e 1. Utilizzando la formula delle probabilità totali abbiamo

$$P(X_i = 1) = \sum_{j=1}^{10} P(a_i = j, a_{i+1} = j) = 10 \cdot \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{10}.$$

Abbiamo quindi che sia X_1 che X_2 sono variabili $B(1, \frac{1}{10})$. In alternativa si può semplicemente osservare che su 100 codici di 2 cifre esattamente 10 hanno le due cifre uguali e quindi la probabilità di avere le due cifre uguali è proprio 1/10.

(b) Determiniamo la densità congiunta $p(x, y)$ di X_1 e X_2. Abbiamo

$$\begin{aligned}
&p(0,0) = P(X_1 = 0, X_2 = 0) = P(a_2 \neq a_1, a_2 \neq a_3) = \frac{10 \cdot 9 \cdot 9}{1000} = \frac{81}{100} \\
&p(0,1) = P(X_1 = 0, X_2 = 1) = P(a_1 \neq a_2 = a_3) = \frac{10 \cdot 9}{1000} = \frac{9}{100} \\
&p(1,0) = p(0,1) \\
&p(1,1)P(X_1 = 1, X_2 = 1) = P(a_1 = a_2 = a_3) = \frac{10}{1000} = \frac{1}{100}
\end{aligned}$$

In tutti i casi abbiamo $p(x, y) = P(X_1 = x)P(X_2 = y)$ e quindi le variabili X_1 e X_2 sono indipendenti.

(c) Sia $Y = X_1 + X_2$. La variabile Y può assumere i valori 0,1,2. Calcoliamo la densità.

$$\begin{aligned}
&P(Y = 0) = P(X_1 = 0, X_2 = 0) = P(X_1 = 0)P(X_2 = 0) = \frac{81}{100} \\
&P(Y = 2) = P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) = \frac{1}{100} \\
&P(Y = 1) = 1 - P(Y = 0) - P(Y = 2) = \frac{18}{100}
\end{aligned}$$

e quindi riassumento

$$p_Y(k) = \begin{cases} \frac{81}{100} & k = 0 \\ \frac{18}{100} & k = 1 \\ \frac{1}{100} & k = 2 \\ 0 & \text{altrimenti.} \end{cases}$$

Poniamo inoltre $Z = X_1X_2$. La variabile Z può assumere solo i valori 0 e 1 ed è quindi una variabile di Bernoulli. Si ha

$$P(Z = 1) = P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) = \frac{1}{100}$$

e quindi $Z \sim B(1, \frac{1}{100})$.

(2) Si considerino 2 variabili aleatorie indipendenti X ed Y. Sono noti i seguenti valori della densità congiunta

$X \backslash Y$	0	1
0	$\frac{1}{8}$	$\frac{3}{8}$
1		$\frac{3}{32}$
2		

(a) Calcolare $P(X = 0)$.
(b) Calcolare $P(Y = 1)$.
(c) Completare la tabella a doppia entrata della densità congiunta delle variabili X ed Y.
(a) Dalla prima riga della tabella deduciamo che

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}.$$

(b) Siccome le variabili X ed Y sono indipendenti abbiamo

$$P(X = 0, Y = 1) = P(X = 0)P(Y = 1).$$

Dalla tabella sappiamo che $P(X = 0, Y = 1) = \frac{3}{8}$ mentre dal punto (a) abbiamo $P(X = 0) = \frac{1}{2}$. Ne deduciamo che

$$P(Y = 1) = \frac{P(X = 0, Y = 1)}{P(X = 0)} = \frac{3}{4}.$$

(c) Analogamente a prima dall'identità $P(X = 1, Y = 1) = P(X = 1)P(Y = 1)$ possiamo determinare

$$P(X = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{\frac{3}{32}}{\frac{3}{4}} = \frac{1}{8}.$$

Possiamo quindi completare la descrizione delle densità marginali: $P(X = 2) = 1 - P(X = 0) - P(X = 1) = \frac{3}{8}$ e $P(Y = 0) = 1 - P(Y = 1) = \frac{1}{4}$. I valori mancanti della densità congiunta sono quindi

$$P(X = 1, Y = 0) = P(X = 1)P(Y = 0) = \frac{1}{8} \cdot \frac{1}{4} = \frac{1}{32}$$
$$P(X = 2, Y = 0) = P(X = 2)P(Y = 0) = \frac{3}{8} \cdot \frac{1}{4} = \frac{3}{32}$$
$$P(X = 2, Y = 1) = P(X = 2)P(Y = 1) = \frac{3}{8} \cdot \frac{3}{4} = \frac{9}{32}$$

(3) Una parola della lingua italiana è composta da un numero di lettere di media 5 e scarto quadratico medio 2. Si consideri una frase casuale composta da 20 parole casuali e sia X il numero di caratteri (spazi inclusi) di cui è composta una tale frase.
(a) Determinare la media e la varianza di X.
(b) Determinare la probabilità con cui una frase casuale di 20 parole è costituita da più di 100 caratteri.
(a) Detta X_i la variabile lunghezza della i-esima parole abbiamo

$$X = X_1 + X_2 + \cdots + X_{20} + 19,$$

in quanto tra le 20 parole dobbiamo considerare anche i 19 spazi. Si ha quindi

$$E[X] = E[X_1] + \cdots + E[X_{20}] + E[19] = 20 \cdot 5 + 19 = 119.$$

e

$$Var(X) = Var(X_1) + \cdots + Var(X_{20}) = 20 \cdot 4 = 80$$

in quanto la varianza di una variabile costante è nulla.

(b) La variabile X la approssimiamo con una variabile $\zeta \sim N(119, 80)$ grazie al teorema centrale del limite. Applicando la correzione di continuità abbiamo

$$P(X > 100) = P(\zeta > 100,5) = P(\zeta_0 > \frac{100,5 - 119}{\sqrt{80}}) = P(\zeta_0 > -2,07) = \Phi(2,07) = 98\%.$$

26/02/2013

(1) Un'urna contiene 6 palline, numerate da 1 a 6. Le palline sono anche state colorate, 4 rosse e 2 blu, in modo del tutto casuale, e non sappiamo come.
   (a) Estraendo 3 palline, determinare la probabilità che esattamente 2 di esse siano rosse;
   (b) estraendo 3 palline, determinare la probabilità di estrarre una pallina rossa numerata con il 6 (cioè che la pallina numerata con il 6 sia stata colorata di rosso ed estratta);
   (c) determinare la probabilità che le palline numerate con 1,4 e 6 siano tutte rosse;
   (d) estraendo 3 palline, determinare la probabilità che la somma dei numeri che compaiono sulle palline rosse estratte sia 11.

(2) Un paese in Giappone si trova vicino ad un vulcano ed è anche a rischio terremoto. Si sa che il vulcano erutta mediamente ogni 10 anni e che si verifica un terremoto mediamente ogni 15 anni. Si può assumere che i tempi di attesa per entrambi i fenomeni siano variabili esponenziali, indipendenti tra loro.
   (a) Determinare la probabilità che si verifichi un terremoto nei prossimi 10 anni;
   (b) Dovendo andare a vivere in questo paese per i prossimi 5 anni, stabilire la probabilità che in questo periodo si verifichi un'eruzione e/o un terremoto, sapendo che negli ultimi 10 anni non si sono verificati né eruzioni, né terremoti.

(3) Viene misurata l'altezza in centimetri in un campione di 101 watussi ottenendo i seguenti risultati :

$170-175$	6
$175-180$	11
$180-185$	13
$185-190$	20
$190-195$	1
$195-200$	40
$200-210$	10

   (a) Determinare media, varianza e classe mediana e classe modale del campione;
   (b) Determinare l'intervallo di confidenza al 95% per la media della variabile aleatoria "altezza di watusso scelto a caso" in base ai dati sperimentali.

13 Giugno 2013

(1) Un'urna contiene 3 palline rosse e 2 palline bianche. Le estraiamo senza rimpiazzo finché non otteniamo almeno una pallina per ciascun colore. Denotiamo con X il numero di palline bianche estratte e Y il numero di palline rosse estratte. Denotiamo inoltre con $Z = X + Y$ il numero totale di palline estratte.

(a) Determinare $P(Z = 2)$;
Si ha $Z = 2$ se la prima è bianca e la seconda rossa o viceversa. Detto B_i l'evento "i.esima estratta bianca" e similmente R_i abbiamo quindi

$$P(Z = 2) = P(B_1 \cap R_2) + P(R_1 \cap B_2) = P(B_1)P(R_2|B_1) = P(R_1)P(B_2|R_1) = \frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{5};$$

(b) determinare la densità di Z; Zpuò assumere solo i valori 2,3,4. Basterà quindi determinare

$$P(Z = 4) = P(R_1 \cap R_2 \cap R_3 \cap B_4) = \frac{3}{5}\frac{2}{4}\frac{1}{3} = \frac{1}{10}$$

in quanto avremo poi $P(Z = 3) = 1 - P(Z_2) - P(Z = 4) = \frac{3}{10}$. La densità di Z è quindi

$$p_Z(k) = \begin{cases} \frac{3}{5} & k = 2 \\ \frac{3}{10} & k = 3 \\ \frac{1}{10} & k = 4 \\ 0 & \text{altrimenti.} \end{cases}$$

(c) determinare $E[Z]$;
Basta utilizzare la definizione

$$E[Z] = \sum k p_Z(k) = 2 \cdot \frac{3}{5} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{1}{10} = \frac{5}{2}.$$

(d) stabilire se X ed Y sono indipendenti;
X ed Y sono dipendenti in quanto, ad esempio, $Y = 3$ implica $X = 1$.

(2) Si considerino due variabili X ed Y indipendenti che assumono valori 1,2,3 definite su uno spazio di probabilità Ω. Sono noti i seguenti valori della loro densità congiunta $p(x, y)$:

$X \backslash Y$	1	2	3
1	$\frac{1}{12}$	$\frac{1}{18}$	
2		$\frac{1}{6}$	$\frac{1}{12}$
3	$\frac{1}{6}$		

(a) Sfruttando l'indipendenza di X ed Y, verificare che $p(1, 2) \cdot p(2, 1) = p(1, 1) \cdot p(2, 2)$ ed utilizzare questa relazione per determinare $p(1, 2)$;
Deriva semplicemente dal fatto che X ed Y sono indipendenti. Abbiamo infatti $p(1, 2) = p_X(1)p_Y(2)$ e similmente le altre da cui $p(1, 2)p(2, 1) = p_X(1)p_Y(2)p_X(2)p_Y(1)$ e $p(1, 1) \cdot p(2, 2) = p_X(1)p_Y(1)p_X(2)p_Y(2)$. Ne segue

$$p(1, 2) = p(1, 1) \cdot p(2, 2)/p(2, 1) = \frac{1}{12} \cdot \frac{1}{6} \cdot 18 = \frac{1}{4}.$$

(b) calcolare $p_X(1)$
$p_X(1)$ lo otteniamo sommando gli elementi nella prima colonna: $p_X(1) = \frac{1}{12} + \frac{1}{4} + \frac{1}{6} = \frac{1}{2}$;

(c) determinare la densità marginale di Y e la densità marginale di X
Dalla prima colonna otteniamo $P(X=1,Y=1)=P(X=1)P(Y=1)$ da cui $P(Y=1)=\frac{1}{12}/\frac{1}{2}=\frac{1}{6}$ e similmente $P(X=1,Y=2)=P(X=1)P(Y=2)$ da cui $P(Y=2)=\frac{1}{4}/\frac{1}{2}=\frac{1}{2}$; infine da $P(X=1,Y=3)=P(X=1)P(Y=3)$ da cui $P(Y=3)=\frac{1}{6}/\frac{1}{2}=\frac{1}{3}$. La densità marginale della X la otteniamo similmente ragionando sulla seconda riga ottenendo $P(X=1)=\frac{1}{2}$, $P(X=2)=\frac{1}{3}$ e $P(X=3)=\frac{1}{6}$

(d) completare la tabella della densità congiunta $p(x,y)$;
È sufficiente a questo punto moltiplicare le densità marginali per ottenere la congiunta.

(e) completare la tabella in modo che le variabili X ed Y siano dipendenti.
Basta compilare la tabella con numeri non negativi diversi dai precedenti in modo che la somma di tutti i coefficienti sia 1. Si possono ad esempio mettere tutti 0 tranne in un posto in cui inserire 1 meno la somma dei numeri che compaiono; o anche, senza fare conti, scambiando due dei valori che si erano ottenuti nel caso indipendente.

(3) Si consideri la seguente funzione dipendente da un parametro reale a

$$f(x)=\begin{cases} a(1-x^2) & \text{se } x\in[-1,1] \\ 0 & \text{altrimenti.}\end{cases}$$

(a) Stabilire per quale valore di a la funzione $f(t)$ è la densità di una variabile aleatoria continua X;
Siccome $\int_{-\infty}^{+\infty} f(x)\,dx=\frac{4}{3}a$ deduciamo che necessariamente $a=\frac{3}{4}$;

(b) deteminare $E[X]$;
$E[X]=0$ in quanto la densità è una funzione pari.

(c) determinare $P(X<\frac{2}{3}|X>\frac{1}{3})$.
Abbiamo, per definizione di probabiltià condizionale

$$P(X<\frac{2}{3}|X>\frac{1}{3})=\frac{P(\frac{1}{3}<X<\frac{2}{3})}{P(X>\frac{1}{3})}=\frac{\int_{1/3}^{2/3}\frac{3}{4}(1-x^2)\,dx}{\int_{1/3}^{1}\frac{3}{4}(1-x^2)\,dx}=\frac{5}{7}.$$

01/07/2013

(1) Un dado viene lanciato 420 volte. Poniamo X_i = "numero di lanci che hanno dato come risultato i", per $i = 1, \dots, 6$. Sia inoltre X la somma dei risultati ottenuti in tutti i lanci.
   (a) Determinare la densitá delle X_i;
   (b) determinare la media di X;
   (c) verificare che $\mathrm{Var}(X) = 1225$;
   (d) determinare $P(X \geq 1500)$;
   (e) stabilire se X_1 ed X sono indipendenti;
   (f) detta P' la probabilità condizionale difinita da $P'(A) = P(A|X = 421)$ per ogni evento A, determinare la densità della variabile X_1 rispetto alla probabilità P'.

(2) Siano X e Y due variabili indipendenti uniformi nell'intevallo $[1, 2]$. Poniamo $Z = \min(X, Y)$ e $W = \max(X, Y)$.
   (a) Determinare la funzione di ripartizione e la densità di Z e di W;
   (b) giustificare l'identiã $X + Y = Z + W$;
   (c) determinare $E[Z]$ e $E[W]$;
   (d) stabilire se Z e W sono indipendenti.

(3) Si considerino 30 variabili indipendenti $X_1, \dots, X_{30}$ aventi la seguente funzione di densità

$$f(t) = \begin{cases} 4(t - t^3) & \text{se } 0 \leq t \leq 1 \\ 0 & \text{altrimenti.} \end{cases}$$

   (a) Verificare che $f(t)$ è effettivamente una densità continua;
   (b) deteminare $\mathrm{Var}(X_1)$;
   (c) deteminare $P(X_1 + \cdots + X_{30} > 15)$.

15/07/2013

(1) Un'urna contiene una pallina bianca, una rossa e una gialla. Ne estraiamo una e la reinseriamo insieme ad un'altra pallina dello stesso colore di quella estratta. Estraiamo a questo punto un'altra pallina. Denotiamo con X il numero di palline bianche estratte e con A l'evento "le due palline estratte sono di diverso colore".

(a) Determinare la probabilità di non estrarre palline bianche.
(b) Determinare la densità di X.
(c) Stabilire se gli eventi A e $\{X = 0\}$ sono dipendenti.
(d) Descrivere uno spazio di probabilità che modellizzi questo fenomeno aleatorio.

Soluzione.

(a) Detti R_i, B_i, G_i gli eventi "pallina rossa/bianca/gialla all'estrazione i" abbiamo

$$P(X=0) = P(R_1 \cap R_2) + P(G_1 \cap G_2) + P(R_1 \cap G_2) + P(G_1 \cap G_2) = \frac{1}{6} + \frac{1}{6} + \frac{1}{12} + \frac{1}{12} = \frac{1}{2}.$$

(b) Si ha

$$P(X=2) = P(B_1 \cap B_2) = \frac{1}{6}$$

e quindi $P(X = 1) = 1 - P(X = 0) - P(X = 1) = \frac{1}{3}$. La densità di X è quindi

$$p_X(k) = \begin{cases} \frac{1}{2} & \text{se } k = 0; \\ \frac{1}{3} & \text{se } k = 1 \\ \frac{1}{6} & \text{se } k = 2 \\ 0 & \text{altrimenti.} \end{cases}$$

(c) Abbiamo

$$P(A) = P(R_1 \cap B_2) + P(R_1 \cap G_2) + \cdot + P(G_1 \cap B_2) = 6 \cdot \frac{1}{12} = \frac{1}{2}.$$

e $P(A \cap \{X = 0\}) = P(R_1 \cap G_2) + P(G_1 \cap R_2) = \frac{1}{6}$. Di conseguenza $\frac{1}{6} = P(A \cap \{X = 0\}) \neq P(A)P(X = 0) = \frac{1}{2} \cdot \frac{1}{2}$ e quindi gli eventi A e $\{X = 0\}$ sono dipendenti.

(d) Lo spazio Ω può essere descritto come tutti i possibili risultati delle due estrazioni cioè

$$\Omega = \{bb, rr, gg, br, rb, gr, rg, bg, gb\}.$$

Gli eventi sono dati da tutti i sottoinsiemi di Ω e la probabilità è definita da

$$P(bb) = P(rr) = P(gg) = \frac{1}{6}$$

e

$$P(br) = P(rb) = P(gr) = P(rg) = P(bg) = P(gb) = \frac{1}{12}.$$

(2) Sia X una variabile aleatoria avente funzione di ripartizione

$$F_X(t) = \begin{cases} 0 & \text{se } t < 0 \\ 1 - e^{-t-1} & \text{se } t \geq 0 \end{cases}$$

(a) Disegnare il grafico di $F_X(t)$. $F_X(t)$ è continua?
(b) (*) Osservare che $(-1, 0] = \cup_{n=1}^{\infty} (-\frac{1}{n}, -\frac{1}{n+1}] \cup \{0\}$ e dedurre che $P(X = 0) = 1 - e^{-1}$;
(c) Determinare $P(X = 1)$;
(d) Determinare $P(X > 2 | X > 1)$;
(e) Determinare $P(X > 1 | X > 2)$.

Soluzione

(a) Il grafico di $F_X(t)$ è costante uguale a 0 sulla semiretta $\mathbb{R}_{<0}$, vale $1-e^{-1}$ in 0 e poi tende ad 1 per $t \to +\infty$. La funzione $F_X(t)$ è quindi continua per ogni $t \neq 0$, ma è discontinua in $t = 0$.

(b) La prima parte è del tutto evidente scrivendola come

$$(-1,0] = (-1,-\frac{1}{2}] \cup (-\frac{1}{2},-\frac{1}{3}] \cup \cdots \cup \{0\}.$$

Siccome gli intervalli nella destra sono disgiunti abbiamo

$$P(X \in (-1,0]) = P(X \in (-1,-\frac{1}{2}]) + P(X \in (-\frac{1}{2},-\frac{1}{3}]) + \cdots + P(\{X=0\})$$

da cui, ricordando che $P(X \in (a,b]) = F_X(b) - F_X(a)$ per ogni a, b otteniamo $P(\{X = 0\}) = F_X(0) = 1 - e^{-1}$.

(c) Siccome la funzione di ripartizione è continua in 1 abbiamo $P(X = 1) = 0$.

(d) Per definizione di probabilità condizionale abbiamo

$$P(X > 2|X > 1) = \frac{P(X > 2, X > 1)}{P(X > 1)} = \frac{P(X > 2}{P(X > 1)} = \frac{1 - F_X(2)}{1 - F_X(1)} = \frac{e^{-3}}{e^{-2}} = e^{-1}.$$

(e) $P(X > 1|X > 2) = 1$, in quanro, chiaramente. se $X > 2$ abbiamo anche $X > 1$.

(3) I voti ottenuti da uno studente sono 20,24,21,25,25,20,21,21,20,26.

(a) Determinare le modalità e le rispettive frequenze;
(b) Determinare la media e la varianza;
(c) Determinare la mediana e la moda;
(d) Determinare la media armonica e la media geometrica.

Soluzione.

(a) Le modalità sono 20,21,24,25,26 e le rispettive frequenze sono 3,3,1,2,1.

(b) Detto x il carattere in esame la sua media è

$$\mu_x = \frac{1}{10}(3 \cdot 20 + 3 \cdot 21 + 24 + 2 \cdot 25 + 26) = 22.3.$$

Per calcolare la varianza determiniamo prima la media di x^2

$$\mu_{x^2} = \frac{1}{10}(3 \cdot 20^2 + 3 \cdot 21^2 + 24^2 + 2 \cdot 25^2 + 26^2) = 502.5$$

da cui la varianza di x è

$$\sigma_x^2 = \mu_{x^2} - \mu_x^2 = 5.21.$$

(c) La mediana si ottiene riordinando i valori 20,20,20,21,21,21,24,25,25,26 per poi scegliere la media aritmetica dei due valori centrali (dato che il campione ha cardinalità pari). In questo caso i due valori centrali sono entrambi 21 e quindi la mediana è 21.
Il carattere risulta essere bimodale, e le due mode sono 21 e 20.

(d) La media armonica si ottiene facendo l'inverso della media degli inversi:

$$\Big(\frac{1}{10}\Big(3 \cdot \frac{1}{20} + 3 \cdot \frac{1}{21} + \frac{1}{24} + 2 \cdot \frac{1}{25} + \frac{1}{26}\Big)\Big)^{-1} = 22.08$$

La media geometrica è

$$(20^3 \cdot 21^3 \cdot 24 \cdot 25^2 \cdot 26)^{1/10} = 22.19$$

05/09/2013

(1) Un mazzo contiene sette carte numerate da 1 a 7. Ne estraiamo due a caso senza rimpiazzo. Si considerino gli eventi A ="la prima carta estratta è dispari" e B ="le due carte pescate hanno numeri consecutivi".
   (a) Determinare $P(A)$;
   (b) determinare $P(B)$;
   (c) stabilire se A e B sono indipendenti;
   (d) determinare il coefficiente di correlazione ρ_{χ_A,χ_B}; si ha correlazione diretta o inversa? Darne una spiegazione intuitiva.
   (e) rispondere alla domanda (c) nel caso in cui le due estrazioni venissero effettuate con rimpiazzo.

(2) Si considerino due variabili continue indipendenti X ed Y definite su uno stesso spazio di probabilità Ω. La variabile X è uniforme nell'intervallo $[0,1]$ mentre la variabile Y è normale di media $1/2$ e varianza $1/4$.
   (a) Determinare $P(X > 1/3)$ e $P(Y > 1/3)$;
   (b) determinare la probabilità che almeno una delle due variabili X ed Y sia maggiore di $1/3$.
   (c) determinare la funzione di ripartizione di X^2.

(3) I voti ottenuti da uno studente nei compiti di matematica nella sua carriera sono riportati nella seguente tabella:

Voto	Frequenza
4	2
5	3
6	7
7	4
8	2
9	1

   (a) Determinare media e varianza del campione in esame;
   (b) detta X la variabile "voto al prossimo compito" determinare l'intervallo di confidenza al 95% per la media di X;
   (c) determinare l'intervallo di confidenza al 95% per la probabilità che il prossimo compito sarà sufficiente.

13/01/2014

(1) Un mazzo contiene dieci carte numerate da 1 a 10. Lanciamo un dado e peschiamo un numero di carte pari al numero uscito sul dado. Denotiamo con X il numero uscito sul dado e con Y la variabile di Bernoulli che vale 1 se e solo se la somma delle carte estratte è dispari.
   (a) Determinare la densità di X;
   (b) determinare $P(X \text{ è dispari }, Y = 1)$;
   (c) determinare $P(Y = 1|X = 2)$;
   (d) determinare la densità di Y;
   (e) stabilire se X ed Y sono indipendenti.

(2) Siano X ed Y due variabili aleatorie indipendenti definite su uno stesso spazio di probabilità Ω. Le variabili X ed Y sono entrambe esponenziali di media 1/2..
   (a) Determinare $P(X > 1/3)$;
   (b) determinare la densità della variabile $Z = 2X + 1$;
   (c) determinare la probabilità che Y e Z siano entrambe maggiori di 2;
   (d) determinare la probabilità che X e Z siano entrambe maggiori di 2.

(3) La produzione annuale di una pianta di peperoni è una variabile normale di media 5kg e scarto quadratico medio 700g
   (a) Stabilire la probabilità che una piantina produca almeno 5.5kg di peperoni;
   (b) Stabilire la probabilità che un campo di 120 piante produca almeno 700kg di peperoni;

17/02/2014

(1) In un cinema con tre sale vengono proiettati i film A, B, C. Il film A viene scelto dal 30% degli spettatori, il film B dal 50% e il film C dal 20%. Da un'indagine statistica si sa che il film A è piaciuto all'80% delle persone che lo hanno visto, il B al 70% e il C al 60%.
   (a) Qual è la probabilità che a una persona che esce da questo cinema sia piaciuto il film che ha visto?
   (b) Su 300 persone che escono dal cinema, qual è la probabilità che almeno a 250 persone sia piaciuto il film visto?
   (c) Su 6 persone cui è piaciuto il film, qual è la probabilità che almeno 4 di esse abbiano visto il film B?

(2) Un dado rosso, uno blu e uno verde vengono lanciati simultaneamente per 4 volte. Poniamo X la somma dei valori ottenuti dai dadi rossi, Y la somma dei valori ottenuti dai dadi rossi e blu, e Z la somma di tutti i valori ottenuti.
   (a) Determinare $E[X]$.
   (b) Stabilire se X, Y e Z sono a due a due dipendenti.
   (c) Determinare la media di $X + Y + Z$.
   (d) Determinare $P(Y - X \geq 5)$

(3) Sia X una variabile continua e sia

$$f_X(t) = \begin{cases} \frac{t}{2} & \text{se } 0 \leq t \leq 2] \\ 0 & \text{altrimenti} \end{cases}$$

la sua densità.
   (a) Determinare $E[X]$ e $E[X^2]$.
   (b) Se $X_1, \ldots, X_{81}$ sono 81 variabili indipendenti aventi la stessa densità di X determinare

$$P(X_1 + \cdots + X_{81}) > 100.$$