

2. (2 punti) Si scriva la tabella delle frequenze cumulate per il carattere SATELLITI e si rappresenti graficamente la risultante funzione cumulativa delle frequenze.



3. (3 punti) Si scriva la distribuzione di frequenza per il carattere VELOCITA', riclassificato nelle seguenti classi: [0,10), [10,30), [30,50). Tale carattere è più o meno variabile del carattere SATELLITI? Si giustifichi la risposta operando sui dati riclassificati.

V	n_i	p_i	
[0,10)	4	0.4444	Per confrontare la variabilità dei due caratteri si utilizza il coefficiente di variazione. $M(V) = [(5 \times 4) + (20 \times 3) + (40 \times 2)] / 9 = 17.7778$ $M(V^2) = [(25 \times 4) + (400 \times 3) + (1600 \times 2)] / 9 = 500$ $Var(V) = M(V^2) - M^2(V) = 183.9498$ $CV(V) = SQM(V) / M(V) = 13.5628 / 17.7778 = 0.7629$ $M(S) = 6.4444$ $M(S^2) = 109.5556$ $Var(S) = M(S^2) - M^2(S) = 68.0253$ $CV(S) = 8.2478 / 6.4444 = 1.2798$ $CV(S) > CV(V)$ quindi S è più variabile.
[10,30)	3	0.3333	
[30,50)	2	0.2222	

4. (2 punti) Si enunci la disuguaglianza di Chebyshev, specificando tutte le quantità che compaiono nella formula. Tale disuguaglianza dipende dalla forma della distribuzione dei dati? Si giustifichi la risposta.

Per un carattere X con media μ e scarto quadratico medio σ vale

$$Fr\{ \mu - \lambda\sigma < X < \mu + \lambda\sigma \} \geq 1 - 1/\lambda^2$$

con $\lambda > 1$.

La disuguaglianza vale qualunque sia la forma della distribuzione, poichè basta conoscere media e scarto quadratico medio per calcolarla.

5. (2 punti)

a) Si costruisca la tabella a doppia entrata per i caratteri DIMENSIONE e VELOCITA', quest'ultimo riclassificato come nell'esercizio 3.

	V	[0,10)	[10,30)	[30,50)	
DIM					
1		0.1111 (1)	0.2222 (2)	0.2222 (2)	0.5555 (5)
2		0.3333 (3)	0.1111 (1)	0 (0)	0.4444 (4)
		0.4444 (4)	0.3333 (3)	0.2222 (2)	

La tabella è costruita con le frequenze relative.
In parentesi sono indicate le frequenze assolute.

b) Si costruisca la distribuzione di frequenza di VELOCITA' subordinata a DIMENSIONE=1.

V	P(V D=1)
[0,10)	$0.1111/0.5555 = 0.2$
[10,30)	$0.2222/0.5555 = 0.4$
[30,50)	$0.2222/0.5555 = 0.4$

6. (2 punti) Siano X e Y due caratteri qualsiasi. Si dica se le seguenti affermazioni sono vere o false

- Se X è statisticamente indipendente da Y, allora $Cov(X,Y)=0$
- Se X è regressivamente indipendente da Y, allora Y è regressivamente indipendente da X
- Se $Cov(X,Y)=0$, allora X è statisticamente indipendente da Y
- Se X è statisticamente indipendente da Y, allora X è regressivamente indipendente da Y e Y è regressivamente indipendente da X

a) VERO
b) FALSO
c) FALSO
d) VERO

7. (4 punti) E' stato scoperto un nuovo pianeta nel sistema solare, che dista dal sole 650 miliardi di km ed ha una velocità orbitale di 3 km/s. Se ne vuole prevedere il periodo di rivoluzione.

- (3 punti) Sapendo che il coefficiente di correlazione lineare tra RIVOLUZIONE e DISTANZA è pari a -0.1807, si stabilisca, calcolando un opportuno indice, quale tra i caratteri DISTANZA e VELOCITA' è migliore ai fini della previsione. (Il carattere VELOCITA' è da intendersi *non*

riclassificato in intervalli. Per il calcolo di media e varianza si possono però utilizzare i risultati dell'esercizio 3).

Il carattere migliore ai fini della previsione è quello che ha un legame lineare più forte con il carattere che si vuole prevedere. E' pertanto necessario confrontare i valori assoluti dei due coefficienti di correlazione lineare.

$$Var(R) = M(R^2) - M^2(R) = 84481.4444 - 44708.7343 = 39772.7101$$

$$Var(V) = M(V^2) - M^2(V) = 500 - 316.0502 = 183.9498$$

$$Cov(R,V) = M(R \times V) - M(R)M(V) = 4731.9222 - 3759.0163 = 972.9059$$

$$\rho(R,V) = Cov(R,V) / \sqrt{SQM(R) SQM(V)} = 972.9059 / 2704.8428 = 0.3597$$

$|\rho(R,V)| > |\rho(R,DIS)|$, pertanto V è da preferire.

b) (1 punto) Sulla base di quanto stabilito al punto (a), si preveda il valore del carattere RIVOLUZIONE per il nuovo pianeta.

I coefficienti della retta di regressione sono

$$b = Cov(V,R) / Var(V) = 5.289$$

$$a = M(R) - b M(V) = 117.4171$$

Il valore previsto per R, in corrispondenza di V=3, è

$$R = a + 3b = 133.2841$$

8. (3 punti) Da un'urna contenente 3 palline bianche e 7 nere si estraggono due palline senza reimmissione.

a) (1 punto) Si calcoli la probabilità che entrambe le palline estratte siano bianche.

B_1 =la prima pallina estratta è bianca, B_2 =la seconda pallina estratta è bianca, e similmente per le palline nere si hanno N_1 e N_2 .

$$P(\text{entrambe le palline sono bianche}) = P(B_1 \cap B_2) = 3/10 \times 2/9 = 0.0667$$

b) (1 punto) Sapendo che la prima pallina estratta è bianca, si calcoli la probabilità che esca di nuovo una pallina bianca alla seconda estrazione.

Si utilizza la formula per il calcolo della probabilità condizionale

$$P(B_2 | B_1) = P(B_1 \cap B_2) / P(B_1) = 1/15 \times 10/3 = 0.2222$$

c) (1 punto) Si supponga che le estrazioni avvengano nel modo seguente: se la prima pallina estratta è bianca, si procede alla seconda estrazione senza reimmissione; se la prima pallina estratta è nera, la si rimette nell'urna, se ne aggiunge un'altra nera e si procede alla seconda estrazione. Si calcoli la probabilità che la seconda pallina estratta sia nera.

Si utilizza il teorema delle probabilità totali

$$\begin{aligned} P(N_2) &= P(N_2 | N_1)P(N_1) + P(N_2 | B_1)P(B_1) \\ &= 7/10 \times 8/11 + 3/10 \times 7/9 \\ &= 0.7424 \end{aligned}$$

9. (3 punti) Il signor Rossi per recarsi al lavoro deve prendere un treno e un autobus. Per non arrivare tardi al lavoro non deve perdere né il treno né l'autobus. Se perde il treno perde di sicuro anche l'autobus (Si ipotizzi pari a 0.5 la probabilità di prendere il singolo mezzo e la probabilità di prendere l'autobus dato che si è preso il treno).

a) (1 punto) Qual è la probabilità che il signor Rossi arrivi puntuale al lavoro?

Sia T = "il signor Rossi prende il treno", \bar{T} = "il signor Rossi non prende il treno" e similmente per l'autobus A e \bar{A} . Sia E l'evento "il signor Rossi arriva puntuale": questo si verifica se e solo se il signor Rossi prende sia il treno sia l'autobus.

$$P(E) = P(T \cap A) = P(T) \cdot P(A | T) = 0.5 \cdot 0.5 = 0.25$$

b) (2 punti) Si calcoli la probabilità che il signor Rossi arrivi in orario almeno 2 volte durante la settimana lavorativa (lunedì-venerdì).

Sia X = numero di giorni in cui il signor Rossi arriva puntuale durante la settimana. Allora X è una binomiale di parametri 5 e 0.25.

$$P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - (0.75^5 + 5 \times 0.25 \times 0.75^4) = 1 - (0.2373 + 0.3955) = 0.3672$$

2. (3 punti) Si scriva la tabella delle frequenze cumulate per il carattere RIVOLUZIONE riclassificato come nell'esercizio 1. Si calcoli la frequenza relativa con cui la variabile assume valori nell'intervallo (160,350].

R	p_i	$F(x_{i+1})$
[0,50)	0.2222	0.2222
[50,150)	0.2222	0.4444
[150,300)	0.3333	0.7778
[300,700)	0.2222	1

L'equazione della funzione di ripartizione negli intervalli [150,300) e [300,700) è rispettivamente
 $F(x) = 0.4444 + 0.0022(x-150)$
 $F(x) = 0.7778 + 0.0005(x-300)$.

Si vuole calcolare
 $Fr\{160 < R \leq 350\} = F\{350\} - F\{160\} = [0.7778 + 0.0005(350-300)] - [0.4444 + 0.0022(160-150)] =$
 $= 0.7778 + 0.025 - 0.4444 - 0.022 = 0.3364$

3. (3 punti) Si scriva la distribuzione di frequenza per il carattere SATELLITI. Tale carattere è più o meno variabile del carattere RIVOLUZIONE? Si giustifichi la risposta (per il carattere RIVOLUZIONE, si operi sui dati riclassificati).

S	n_i	p_i	Per confrontare la variabilità dei due caratteri si utilizza il coefficiente di variazione.
0	2	0.2222	
1	3	0.3333	$M(S) = 58/9 = 6.4444$
2	1	0.1111	$Var(S) = M(S^2) - M^2(S) = 109.5556 - 41.5303 = 68.0253$
15	2	0.2222	$CV(S) = SQM(S)/ M(S) = 8.2478/6.4444 = 1.2798$
23	1	0.1111	

$M(R) = [(25 \times 2) + (100 \times 2) + (225 \times 3) + (500 \times 2)]/9 = 213.8889$
 $M(R^2) = [(625 \times 2) + (10000 \times 2) + (50625 \times 3) + (250000 \times 2)]/9 = 74791.6667$
 $Var(R) = 29043.2052$
 $CV(R) = 170.4207/213.8889 = 0.7968$

$CV(S) > CV(R)$ quindi S è più variabile.

4. (2 punti) Si enunci la disuguaglianza di Chebyshev, specificando tutte le quantità che compaiono nella formula. Tale disuguaglianza dipende dalla forma della distribuzione dei dati? Si giustifichi la risposta.

<p>Per un carattere X con media μ e scarto quadratico medio σ vale $Fr\{ \mu - \lambda\sigma < X < \mu + \lambda\sigma \} \geq 1 - 1/\lambda^2$ con $\lambda > 1$.</p> <p>La disuguaglianza vale qualunque sia la forma della distribuzione, poichè basta conoscere media e scarto quadratico medio per calcolarla.</p>

5. (2 punti)

a) (1 punto) Si costruisca la tabella a doppia entrata per i caratteri DIMENSIONE e RIVOLUZIONE, quest'ultimo riclassificato come nell'esercizio 1.

DIM \ R	[0,50)	[50,150)	[150,300)	[300,700)	
1	0 (0)	0.1111 (1)	0.2222 (2)	0.2222 (2)	0.5555 (5)
2	0.2222 (2)	0.1111 (1)	0.1111 (1)	0 (0)	0.4444 (4)
	0.2222 (2)	0.2222 (2)	0.3333 (3)	0.2222 (2)	

La tabella è costruita con le frequenze relative. In parentesi sono indicate le frequenze assolute.

b) (1 punto) Si costruisca la distribuzione di frequenza di RIVOLUZIONE subordinata a DIMENSIONE=2.

R	P(R D=2)
[0,50)	$0.2222/0.4444 = 0.5$
[50,150)	$0.1111/0.4444 = 0.25$
[150,300)	$0.1111/0.4444 = 0.25$
[300,700)	0

6. (2 punti) Siano X e Y due caratteri qualsiasi. Si dica se le seguenti affermazioni sono vere o false

- Se X è statisticamente indipendente da Y, allora $Cov(X,Y)=0$
- Se X è regressivamente indipendente da Y, allora Y è regressivamente indipendente da X
- Se $Cov(X,Y)=0$, allora X è statisticamente indipendente da Y
- Se X è statisticamente indipendente da Y, allora X è regressivamente indipendente da Y e Y è regressivamente indipendente da X

- VERO
- FALSO
- FALSO
- VERO

7. (4 punti) E' stato scoperto un nuovo pianeta nel sistema solare, che dista dal sole 50 miliardi di km ed ha un periodo di rivoluzione di 10 giorni. Se ne vuole prevedere la velocità orbitale.

a) (3 punti) Sapendo che il coefficiente di correlazione lineare tra RIVOLUZIONE e VELOCITA' è pari a 0.3597, si stabilisca, calcolando un opportuno indice, quale tra i caratteri DISTANZA e RIVOLUZIONE è migliore ai fini della previsione.

Il carattere migliore ai fini della previsione è quello che ha un legame lineare più forte con il carattere che si vuole prevedere. E' pertanto necessario confrontare i valori assoluti dei due coefficienti di correlazione lineare.

$$Var(DIS) = M(DIS^2) - M^2(DIS) = 73550.0556 - 31886.0663 = 41663.9893$$

$$Var(V) = M(V^2) - M^2(V) = 594.3889 - 384.16 = 210.2289$$

$$Cov(DIS, V) = M(DIS \times V) - M(DIS)M(V) = 1247.2444 - 3499.9073 = -2252.6629$$

$$\rho(DIS, V) = Cov(DIS, V) / SQM(DIS) SQM(V) = -2252.6629 / 2959.5623 = -0.7611$$

$|\rho(R, V)| < |\rho(DIS, V)|$, pertanto DIS è da preferire.

b) (1 punto) Sulla base di quanto stabilito al punto (a), si preveda il valore del carattere VELOCITA' per il nuovo pianeta.

I coefficienti della retta di regressione sono

$$b = Cov(DIS, V) / Var(DIS) = -0.0541$$

$$a = M(V) - b M(DIS) = 29.2605$$

Il valore previsto per V, in corrispondenza di DIS=50, è

$$V = a + 50b = 26.5555$$

8. (3 punti) Da un'urna contenente 9 palline bianche e 3 nere si estraggono due palline senza reimmissione.

a) (1 punto) Si calcoli la probabilità che entrambe le palline estratte siano nere.

N_1 =la prima pallina estratta è nera, N_2 =la seconda pallina estratta è nera, e similmente per le palline bianche si hanno B_1 e B_2 .

$$P(\text{entrambe le palline sono nere}) = P(N_1 \cap N_2) = 3/12 \times 2/11 = 0.0455$$

b) (1 punto) Sapendo che la prima pallina estratta è nera, si calcoli la probabilità che esca di nuovo una pallina nera alla seconda estrazione.

Si utilizza la formula per il calcolo della probabilità condizionale

$$P(N_2 | N_1) = P(N_1 \cap N_2) / P(N_1) = 1/22 \times 12/3 = 0.1818$$

c) (1 punto) Si supponga che le estrazioni avvengano nel modo seguente: se la prima pallina estratta è nera, si procede alla seconda estrazione senza reimmissione; se la prima pallina estratta è bianca, la si rimette nell'urna, se ne aggiunge un'altra bianca e si procede alla seconda estrazione. Si calcoli la probabilità che la seconda pallina estratta sia bianca.

Si utilizza il teorema delle probabilità totali

$$\begin{aligned} P(B_2) &= P(B_2 | B_1)P(B_1) + P(B_2 | N_1)P(N_1) \\ &= 10/13 \times 9/12 + 9/11 \times 3/12 \\ &= 0.7815 \end{aligned}$$

9. (3 punti) Il signor Rossi per recarsi al lavoro deve prendere un treno e un autobus. Per non arrivare tardi al lavoro non deve perdere né il treno né l'autobus. Se perde il treno perde di sicuro anche l'autobus. (Si ipotizzi pari a 0.5 la probabilità di prendere il singolo mezzo e la probabilità di perdere l'autobus dato che si è preso il treno).

a) (1 punto) Qual è la probabilità che il signor Rossi arrivi in ritardo al lavoro?

Sia T = "il signor Rossi prende il treno", \bar{T} = "il signor Rossi non prende il treno" e similmente per l'autobus A e \bar{A} . Sia E l'evento "il signor Rossi arriva in ritardo": questo si verifica se il signor Rossi perde il treno oppure se perde l'autobus dopo aver preso il treno.

$$P(E) = P(\bar{T}) + P(T) \cdot P(\bar{A} | T) = 0.5 + 0.5 \cdot 0.5 = 0.75$$

b) (2 punti) Si calcoli la probabilità che il signor Rossi arrivi in ritardo non più di 1 volta durante la settimana lavorativa (lunedì-venerdì).

Sia X = numero di giorni in cui il signor Rossi arriva in ritardo durante la settimana. Allora X è una binomiale di parametri 5 e 0.75.

$$P(X \leq 1) = P(X = 0) + P(X = 1) = 0.25^5 + 5 \times 0.75 \times 0.25^4 = 0.010 + 0.0146 = 0.0156$$