Matematica II: Calcolo delle Probabilità e Statistica Matematica

ELT A-Z

Docente: dott. F. Zucca

Esercitazione # 1

1 Esercizi Statistica Descrittiva

Esercizio 1 *I gruppi sanguigni di* 12 *persone sono*

B, B, AB, O, A, O, A, A, A, B, A, A.

Si costruisca la tabella delle distribuzioni di frequenza e diagramma a barre.

Esercizio 2 *Un certo macchinario produce lotti di 100 pezzi ciascuno. Il numero di pezzi difettosi in 25 lotti ispezionati è*

1, 5, 3, 1, 3, 2, 2, 1, 2, 5, 3, 0, 1, 4, 3, 7, 1, 3, 1, 7, 2, 1, 2, 4, 8

1. *Costruire le tabelle di distribuzione delle frequenze e l'istogramma. Determinare la media, i quartili, il quantile* 0.8*, la moda, la varianza campionaria, la differenza interquartile ed il range.*

2. *Per sbaglio, un impiegato scambia nell'istogramma le frequenze dell' 1 e del 4. La moda e la media dei dati così modificati cambiano rispetto a quelle dei dati originali?*

Esercizio 3 *Una compagnia di assicurazioni ha rilevato il numero di incidenti nel periodo 1996-2000 relativo a* 25 *assicurati*

0, 1, 0, 2, 5, 0, 1, 4, 3, 2, 0, 1, 0, 5, 2, 0, 0, 6, 1, 1, 0, 3, 1, 2, 2.

1) *Rappresentare i dati con un istogramma.*

2) *Calcolare media e varianza.*

3) *Calcolare i quartili.*

4) *Con quale frequenza non si è dovuto risarcire più di un sinistro?*

Esercizio 4 *I dati di un esperimento vengono raggruppati in* 4 *classi la cui distribuzione di frequenza è Classe[0-2) = 0.1 Classe[2-4) = 0.2 Classe[4-6) = 0.4 Classe[6-8] = 0.3 Stimare:*

1) la media e la varianza;

2) le classi contenenti i quartili [R: $q_{0.25}$ $\in$Cl[2,4), $q_{0.5}$ $\in$Cl[4,6), $q_{0.75}$ $\in$Cl[6,8)].

2 Esercizi preliminari

Esercizio 5 *Per misurare accuratamente dei pesi viene usata una scala digitale. Sia X la variabile aleatoria che indica la misurazione fatta usando questa scala e si considerino i seguenti intervalli di valori di misurazione:*

A : *peso supera i 20 grammi*
B : *peso è inferiore o uguale a 15 grammi*
C : *peso è compreso tra 15 e 24 grammi (estremi esclusi).*

Si conoscono le seguenti probabilità:

$P(X \in A) = 0.5$
$P(X \in B) = 0.3$
$P(X \in C) = 0.6$

a) A e B sono mutuamente disgiunti? B e C? A e C?

b) Descrivere A^c e determinarne la probabilità.

c) Descrivere C^c e determinarne la probabilità.

d) Determinare $P(15 < X \leq 20)$.

Esercizio 6 *Siano A, B e C insiemi a due a due disgiunti con*

$P(X \in A) = 0.2$
$P(X \in B) = 0.3$
$P(X \in C) = 0.5$

Determinare:

a) $P(X \in A^c)$

b) $P(X \in B^c)$

c) $P(X \in C^c)$

d) $P(X \in A \cup B)$

e) $P(X \in A \cup C)$

Esercizio 7 *Sia* X *la durata (in ore) di un laser semiconduttore con le seguenti probabilità:*

$P(X \leq 5000) = 0.05$
$P(5000 < X \leq 7000) = 0.5$
$P(X > 7000) = 0.45$

a) Determinare $P(X \leq 7000)$

b) Determinare $P(X > 5000)$

c) Supponiamo ora che ci siano tre laser indipendenti tutti soddisfacenti le ipotesi precedenti. Calcolare:

  1) la probabilità che tutti e tre i laser funzionino per più di 7000 ore;

  2) la probabilità che tutti e tre i laser funzionino per più di 5000 ore;

  3) nessuno dei tre laser funzioni per più di 7000 ore.

Esercizio 8 *In un gioco televisivo viene messo in palio un 1 milione di euro. Per vincerlo il concorrente dovrà indovinare fra tre buste qual è quella che contiene l'assegno. Il concorrente sceglie a caso una busta; a questo punto il conduttore mostra una delle due buste che sa essere vuota, offrendo al concorrente di cambiare la propria busta con quella rimanente.*
Qual è la probabilità di vincere il premio conservando la prima busta scelta?
Qual è la probabilità di vincere cambiando la busta?
Qual è la probabilità di vincere se gioca a testa e croce fra le due strategie?

Esercizio 9 *Una popolazione nordeuropea ha un'incidenza di AIDS pari allo* 0.05%*. Si utilizza un test che su una persona ammalata è positivo con una probabilità di* 0.999*, mentre su una persona sana risulta positivo con probabilità* 0.002.

1. *Qual è la probabilità che un individuo con test positivo sia effettivamente affetto da AIDS?*

2. *Qual è la probabilità che un individuo con test negativo non sia affetto da AIDS?*

3. *Qual è la probabilità che il test dica la verità?*

3 Variabili continue

Esercizio 10 *Sia* X *una variabile aleatoria con funzione di ripartizione:*

$$F(x) = \begin{cases} 0, & se\ x < 0 \\ \frac{1}{50}x^2, & se\ 0 \leq x < 5 \\ -\frac{1}{50}x^2 + \frac{2}{5}x - 1, & se\ 5 \leq x < 10 \\ 1, & se\ x \geq 10 \end{cases}$$

a) Disegnare F. Quali valori può assumere la variabile aleatoria (continua) X?

b) Mostrare che X ha densità e calcolarla.

c) Calcolare il valore atteso di X.

Esercizio 11 *Si consideri la funzione*

$$f(x) = \begin{cases} \frac{c}{x^2}, & se\ x > 1 \\ 0, & se\ x \leq 1 \end{cases}$$

a) Determinare $c \in \mathbb{R}$ tale che f sia una funzione di densità.

b) Esistono il valore atteso e la varianza di X? Se sì, calcolarli.

Esercizio 12 *La funzione di densità del tempo (in ore) di rottura di una componente elettronica sia data da*

$$f(x) = \frac{1}{1000} e^{-\frac{x}{1000}},\ per\ x > 0.$$

a) Determinare la probabilità che tale componente duri più di 3000 ore prima di rompersi.

b) Determinare la probabilità che tale componente si rompa nell'intervallo di tempo tra 1000 e 2000 ore.

c) Determinare la probabilità che tale componente si rompa prima di 1000 ore.

d) Determinare il numero di ore in cui con probabilità pari al 10% il componente si è rotto.

Esercizio 13 *Sia*

$$f(x) = \frac{3}{2} x^2,\ per\ -1 < x < 1.$$

Determinare:

1. $P(X > 0)$
2. $P\left(X > \frac{1}{2}\right)$
3. $P\left(-\frac{1}{2} \leq X \leq \frac{1}{2}\right)$, $P\left(|X| \leq \frac{1}{2}\right)$ *e* $P\left(|X| \geq \frac{1}{2}\right)$
4. $P(X < -2)$
5. $P\left(X < 0\ oppure\ X > -\frac{1}{2}\right)$
6. *il valore* $y \in \mathbb{R}$ *tale che* $P(X > y) = 0.05$

Esercizio 14 *Determinare la costante $k \in \mathbb{R}$ tale per cui le seguenti funzioni siano funzioni di densità. Determinare poi la media e la varianza di X.*

1. $f(x) = kx^2$, *per* $0 < x < 4$
2. $f(x) = k(1 + 2x)$, *per* $0 < x < 2$
3. $f(x) = ke^{-x}$, *per* $x > 0$.

4 Variabili discrete

Esercizio 15 *Un'automobile può essere venduta con una serie di optionals.*
La funzione di massa f del numero di optionals scelti dal cliente è data da:

x	7	8	9	10	11	12	13
$f(x)$	0.040	0.130	0.190	0.300	0.240	0.050	0.050

1. *Determinare la probabilità che un cliente scelga meno di 9 optionals.*
2. *Determinare la probabilità che un cliente scelga più di 11 optionals.*
3. *Determinare la probabilità che un cliente scelga un numero di optionals compreso tra 8 e 12 (estremi inclusi).*
4. *Determinare la funzione di ripartizione di X.*
5. *Calcolare il valore atteso e la varianza degli optionals scelti.*

Esercizio 16 *Consideriamo le seguenti funzioni F e G:*

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x < 1 \\ \frac{3}{4}, & 1 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

$$G(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x \leq 1 \\ \frac{3}{4}, & 1 < x < 4 \\ 1, & x \geq 4 \end{cases}$$

1. *Quale delle due è una funzione di ripartizione?*
2. *Risalire a X.*
3. *Calcolare il valore atteso di X.*

Esercizio 17 *Determinare la costante $c \in \mathbb{R}$ tale per cui la seguente funzione è una funzione di massa:*

$$f(x) = cx, \text{ per } x = 1, 2, 3, 4.$$

5 Svolgimenti

Soluzione esercizio 2.

1. La media è

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \frac{1}{n}\sum_{i=1}^{k} \alpha_i n_i \equiv \sum_{i=1}^{k} \alpha_i p_i = 2.88;$$

dove la seconda e terza formula rappresentano, rispettivamente, la media per dati raggruppati in funzione delle frequenze assolute e la media per dati raggruppati in funzione delle frequenze relative. Ricordiamo la formula per il calcolo del quantile $\alpha \in (0,1)$

$$Q(\alpha) := \begin{cases} x_{k+1} & \text{se } k < n\alpha < k+1,\ k \in \mathbb{N} \\ (x_k + x_{k+1})/2 & \text{se } n\alpha = k,\ k \in \mathbb{N}, \end{cases}$$

da cui i quartili risultano $Q_1 := Q(0.25) = 1$, $Q_2 := Q(0.5) = 2$, $Q_3 := Q(0.75) = 4$; il quantile 0.8 è 4.5; la moda 1; la varianza campionaria

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n} x_i^2 - \frac{n}{n-1}\overline{x}^2 \equiv \frac{1}{n-1}\sum_{i=1}^{k} \alpha_i^2 n_i - \frac{n}{n-1}\overline{x}^2 = 4.5267;$$

la differenza interquartile $Q_3 - Q_1 = 3$ ed il range $\max\{x_i : i = 1,\dots,n\} - \min\{x_i : i = 1,\dots,n\} = 8$.

2. Le frequenze assolute dei dati originali sono: $n_0 = 1; n_1 = 7; n_2 = 5; n_3 = 5; n_4 = 2; n_5 = 2; n_7 = 2; n_8 = 1$; mentre quelle dei dati scambiati $n_0^* = 1$; $n_1^* = 2; n_2^* = 5; n_3^* = 5; n_4^* = 7; n_5^* = 2; n_7^* = 2; n_8^* = 1$.

Soluzione esercizio 3.

La media è 1.68; la varianza è 3.1433. Calcoliamo i quartili che risultano 0,1,2. Nel $0.56 \equiv 56\%$ non si è dovuto risarcire più di un sinistro.

Soluzione esercizio 4.

Immaginiamo che la distribuzione sia uniforme all'interno di ogni classe (oppure, in ogni caso, quando non si hanno informazioni più precise, la "migliore" scelta è attribuire ad ogni elemento della classe il valore medio tra gli estremi) pertanto si calcolano: la media 4.8 e la varianza 3.56; le classi contenenti i quartili $q_{0.25} \in$Cl[2,4), $q_{0.5} \in$Cl[4,6), $q_{0.75} \in$Cl[6,8).

Soluzione esercizio 5.

a) $A \cap B = \emptyset$ allora disgiunti

$B \cap C = \emptyset$ allora disgiunti
$A \cap C \neq \emptyset$ allora NON mutuamente disgiunti

b) A^c : peso è inferiore o uguale a 20 grammi
$P(X \in A^c) = 1 - P(X \in A) = 0.5$

c) C^c : peso è ≤ 15 grammi o ≥ 24 grammi
$P(X \in C^c) = 1 - 0.6 = 0.4$

d) $P(15 < X \leq 20) = P(X \leq 20) - P(X \leq 15) = P(X \in A^c) - P(X \in B) = 0.2$

Soluzione esercizio 6.

a) $P(X \in A^c) = 1 - P(X \in A) = 0.8$

b) $P(X \in B^c) = 1 - P(X \in B) = 0.7$

c) $P(X \in C^c) = 1 - P(X \in C) = 0.5$

d) $P(X \in A \cup B) = P(X \in A) + P(X \in B) = 0.5$, siccome A e B sono disgiunti

e) $P(X \in A \cup C) = P(X \in A) + P(X \in C) = 0.7$, siccome A e C sono disgiunti

Soluzione esercizio 7.

a) $P(X \leq 7000) = 1 - P(X > 7000) = 0.55$

b) $P(X > 5000) = 1 - P(X \leq 5000) = 0.95$

c)
1) $P(X_1 > 7000; X_2 > 7000; X_3 > 7000) = (0.45)^3$, siccome i tre laser sono indipendenti
2) $P(X_1 > 5000; X_2 > 5000; X_3 > 5000) = (0.95)^3$, siccome i tre laser sono indipendenti
3) $P(X_1 \leq 7000; X_2 \leq 7000; X_3 \leq 7000) = (0.55)^3$, siccome i tre laser sono indipendenti

Soluzione esercizio 10.

a) I valori assunti (con probabilità positiva) dalla v.a. X sono compresi nell'intervallo $[0, 10]$.

b) F è derivabile con continuità e la funzione di densità $f = F'$ di X risulta essere:

$$f(x) = \begin{cases} \frac{1}{25}x, & 0 < x < 5 \\ -\frac{1}{25}x + \frac{2}{5}, & 5 \leq x < 10 \\ 0, & altrove \end{cases}$$

Quest'ultima è Riemann-integrabile pertanto a norma del Teorema fondamentale del calcolo ammette F come primitiva. (Per una caratterizzazione delle funzioni di ripartizione che ammettono densità è necessario il concetto di funzione assolutamente continua.

c) $E[X] = \int_{\mathbb{R}} x f(x)\, dx = \int_0^5 \frac{1}{25}x^2 dx + \int_5^{10} \left(-\frac{1}{25}x^2 + \frac{2}{5}x\right) dx = 5$

Osservazione Per quanto riguarda l'esistenza di una densità si segnalano alcuni risultati: focalizzate la vostra attenzione sul numero (5). (1) Una funzione è una primitiva (in senso di Lebesgue) di una funzione assolutamente integrabile in (a, b) se e solo se è *assolutamente continua* in (a, b). (2) Una funzione è una primitiva (in senso di Lebesgue) di una funzione assolutamente integrabile in ogni sottointervallo aperto di (a, b) se e solo se è *assolutamente continua* in ogni sottointervallo aperto di (a, b). (3) Se una funzione F è derivabile in (a, b) (tranne al più un numero finito di punti dove però è continua) e la sua derivata è Riemann-integrabile allora F è la primitiva (in senso di Riemann) della sua derivata. (4) Se F è derivabile ovunque tranne al più un insieme finito di punti (dove però è continua) ed F' è Lebesgue-integrabile allora F è la primitiva in senso di Lebesgue di F'. (5) Una funzione di ripartizione che sia derivabile ovunque tranne al più un insieme finito di punti (dove però è continua) ammette densità in senso di Lebesgue (o in senso di Riemann se la derivata è Riemann-integrabile).

Soluzione esercizio 8.

Poichè la probabilità di scegliere la busta contenente la promessa di pagamento è 1/3, se il concorrente decide di conservare la prima busta scelta, la probabilità di vincere è 1/3. Con la seconda strategia –consistente nel cambiare la busta che si ha in mano con la busta rimanente dopo che il conduttore ne ha mostrata una vuota– il concorrente vince se e solo se inizialmente ha scelto una delle due buste vuote. Pertanto, con la strategia del cambio della busta, la probabilità di vincere è pari a 2/3. Poniamo $T = \{\text{Esce testa}\}$, $V = \{\text{Il concorrente vince}\}$ e supponiamo che, se esce testa, il concorrente sceglie la prima strategia, ovvero non cambia la busta. Se gioca a testa o croce fra le due strategie abbiamo, per la formula delle probabilità totali, che:

$$\mathbb{P}(V) = \mathbb{P}(V|T)\mathbb{P}(T) + \mathbb{P}(V|T^c)\mathbb{P}(T^c) = 1/3 * 1/2 + 2/3 * 1/2 = 1/2.$$

Soluzione esercizio 9.

Si considerino gli eventi T_+:="il test è positivo",

T_-:="il test è negativo",
A:="il paziente è ammalato",
S:="il paziente è sano".
Sia $\mathbb{P}(A) = x$, $\mathbb{P}(T_+|A) = p$ e $\mathbb{P}(T_-|S) = q$.

1. Dalla formula di Bayes
$$\mathbb{P}(A|T_+) = \frac{\mathbb{P}(T_+|A)\mathbb{P}(A)}{\mathbb{P}(T_+|A)\mathbb{P}(A) + \mathbb{P}(T_+|S)\mathbb{P}(S)} = \frac{px}{px + q(1-x)}$$
che è crescente in x se $p > q$. Il risultato numerico con $x = 1/2000$, $p = 0.999$ e $q = 0.002$ è $4995/24985$.

2. Analogamente
$$\mathbb{P}(S|T_-) = \frac{\mathbb{P}(T_-|S)\mathbb{P}(S)}{\mathbb{P}(T_-|S)\mathbb{P}(S) + \mathbb{P}(T_-|A)\mathbb{P}(A)} = \frac{(1-q)(1-x)}{(1-q)(1-x) + (1-p)x}$$
Il risultato numerico con $x = 1/2000$, $p = 0.999$ e $q = 0.002$ è $9975010/9975015$.

3. La probabilità cercata è
$$\begin{aligned}\mathbb{P}((T_+ \cap A) \cup (T_- \cap S)) &= \mathbb{P}((T_+ \cap A)) + \mathbb{P}((T_- \cap S)) \\ &= \mathbb{P}((T_+ \cap A))\mathbb{P}(A) + \mathbb{P}((T_- \cap S))\mathbb{P}(S) \\ &= px + (1-q)(1-x)\end{aligned}$$

Il risultato numerico è $19996001/2000000$.

Soluzione esercizio 11.

a) $f \geq 0$ e $\int_{\mathbb{R}} f(x)\, dx = \int_1^{+\infty} \frac{c}{x^2} dx = \left(-\frac{c}{x}\right)\Big|_1^{+\infty} = c$. Di conseguenza, è necessario e sufficiente che $c = 1$.

b) Il valore atteso non è finito.

Soluzione esercizio 12.

a) $P(X > 3000) = \int_{3000}^{+\infty} \frac{1}{1000} e^{-\frac{x}{1000}} dx = \left(-e^{-\frac{x}{1000}}\right)\Big|_{3000}^{+\infty} = e^{-3}$

b) $P(1000 < X < 2000) = \int_{1000}^{2000} \frac{1}{1000} e^{-\frac{x}{1000}} dx = -e^{-2} + e^{-1}$

c) $P(X < 1000) = ... = 1 - \frac{1}{e}$

d) $P(X \leq y) = \int_0^y \frac{1}{1000} e^{-\frac{x}{1000}} dx = -e^{-y/1000} + 1 = 0.1$ allora $y = -1000 \ln(0.9) \approx 105.36$.

   Osserviamo che questo valore è il numero di ore al quale la percentuale attesa di componenti guasti è il 10%; dalla legge forte dei grandi numeri si ha altresì che, se il numero di componenti in uso tende all'infinito, allora il limite della percentuale di quelli guasti tende ($\mathbb{P}$-q.c.) al 10%.

Soluzione esercizio 13.

1. $P(X>0)=\int_0^1 f(x)\,dx=\left(\frac{x^3}{2}\right)\Big|_0^1=\frac{1}{2}$
2. $P\left(X>\frac{1}{2}\right)=\int_{\frac{1}{2}}^1 f(x)\,dx=\left(\frac{x^3}{2}\right)\Big|_{\frac{1}{2}}^1=\frac{7}{16}$
3. $P\left(-\frac{1}{2}\leq X\leq\frac{1}{2}\right)=P\left(|X|\leq\frac{1}{2}\right)=\frac{1}{8}$
   $P\left(|X|\geq\frac{1}{2}\right)=P\left(X\leq-\frac{1}{2}\text{ oppure }X\geq\frac{1}{2}\right)=1-P\left(-\frac{1}{2}\leq X\leq\frac{1}{2}\right)=\frac{7}{8}$

d) $P(X<-2)=0$

4. $P\left(X<0\text{ oppure }X>-\frac{1}{2}\right)=P(X<0)+P\left(X>-\frac{1}{2}\right)-P\left(-\frac{1}{2}<X<0\right)=$
   $=\frac{1}{2}+\frac{9}{16}-\frac{1}{16}=1$
5. y deve risolvere: $1-\frac{y^3}{2}=0.1$. Si ottiene $y=0.965$

Soluzione esercizio 14.

1. $\int_0^4 kx^2dx=k\frac{64}{3}$ allora $k=\frac{3}{64}$
   $E[X]=\int_0^4 kx^3dx=3$
2. $k\int_0^2(1+2x)\,dx=6k$ allora $k=\frac{1}{6}$
   $E[X]=...=\frac{11}{9}$
3. $\int_0^{+\infty}ke^{-x}dx=k$ allora $k=1$
   $E[X]=...=1$

Soluzione esercizio 15.

1. $P(X<9)=0.040+0.0130=0.170$
2. $P(X>11)=0.050+0.050=0.1$
3. $P(8\leq X\leq 12)=0.130+0.190+0.300+0.240+0.050+=0.910$
4.

$$F(x)=\begin{cases}0, & x<7\\ 0.040, & 7\leq x<8\\ 0.170, & 8\leq x<9\\ 0.360, & 9\leq x<10\\ 0.660, & 10\leq x<11\\ 0.900, & 11\leq x<12\\ 0.950, & 12\leq x<13\\ 1 & x\geq 13\end{cases}$$

5. $\mathbb{E}[X] = \sum_i x_i \mathbb{P}_X(x_i) = 9.92$, $\mathrm{var}(X) = \sum_i (x_i - \mathbb{E}[X])^2 p_X(i) \equiv \sum_i x_i^2 p_X(i) - \mathbb{E}[X]^2 = 100.36 - 9.92^2 = 1.9536$.

Soluzione esercizio 16.

1. F è una funzione di ripartizione, mentre G no in quanto non è continua da destra in $x = 1$
2. X assume (con probabilità positiva) i valori 0,1,4
3. $E[X] = \frac{5}{4}$

Soluzione esercizio 17.

c deve esser positiva, inoltre deve soddisfare $1 = c + 2c + 3c + 4c$, quindi per $c = \frac{1}{10}$ la funzione f è di massa.