Probabilità e Statistica
Esercizi

UniVR - Dipartimento di Informatica

Fabio Irimie

2° Semestre 2023/2024

Indice

1 Probabilità elementari e probabilità condizionate

1.1 Esercizio 1

Un corso è frequentato da 10 studenti: 6 maschi e 4 femmine. Viene effettuato un esame ed i punteggi degli studenti sono tutti diversi. Si suppone ciascuna classifica equiprobabile.

1. Qual è la cardinalità dello spazio dei campioni costituito da tutte le possibili classifiche? Qual è una possibile misura di probabilità associata?

$$\Omega \subseteq \mathbb{R}$$

   Tutte le possibili classifiche equivalgono a tutti i modi in cui si possono ordinare i punteggi degli studenti. Quindi, la cardinalità dello spazio degli eventi è 10!:

$$card(\Omega) = 10!$$

   La probabilità associata è calcolata come il rapporto tra il numero di eventi favorevoli e il numero di eventi possibili:

$$P(\omega) = \frac{card(\omega)}{card(\Omega)} = \frac{1}{10!}$$

2. Qual è la probabilità che le quattro studentesse ottengano punteggi migliori?

$$E = \text{"Studentesse ottengono punteggi migliori"}$$

   La probabilità che le studentesse ottengano punteggi migliori è data dal rapporto tra il numero di eventi favorevoli e il numero di eventi possibili:

$$P(E) = \frac{4!}{10!}$$

1.2 Esercizio 2

In uno stock di 100 prodotti, 20 sono difettosi.

1. Dieci vengono scelti a caso, senza rimpiazzo.

   Qual è la probabilità che esattamente la metà siano difettosi?

   Per calcolare la probabilità che esattamente la metà dei prodotti siano difettosi si deve fare il rapporto tra la moltiplicazione del numero di modi in cui si possono scegliere 5 prodotti difettosi e 5 prodotti non difettosi e il numero di modi in cui si possono scegliere 10 prodotti:

$$P(E) = \frac{\binom{20}{5} \cdot \binom{80}{5}}{\binom{100}{10}} = \frac{\frac{20!}{5! \cdot 15!} \cdot \frac{80!}{5! \cdot 75!}}{\frac{100!}{10! \cdot 90!}} = \frac{15504 \cdot 24040016}{1.731030946 \cdot 10^{13}} \approx 0.214 = 21.4\%$$

2. Dieci vengono scelti a caso, con rimpiazzo.

   Qual è la probabilità che esattamente la metà siano difettosi? Si può considerare una variabile di Bernoulli:

$$X_i = \begin{cases} 1 & \text{se il prodotto è difettoso } p = \frac{20}{100} = \frac{1}{5} \\ 0 & \text{altrimenti } p = \frac{80}{100} = \frac{4}{5} \end{cases}$$

La probabilità che esattamente la metà dei prodotti siano difettosi è data dalla distribuzione binomiale:

$$P(\omega) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(E) = \binom{10}{5} \left(\frac{1}{5}\right)^5 \left(\frac{4}{5}\right)^5 = \frac{10!}{5! \cdot 5!} \left(\frac{1}{5}\right)^5 \left(\frac{4}{5}\right)^5 \approx 0.0264 = 2.64\%$$

1.3 Esercizio 3

Consideriamo l'estrazione di una carta da una mazzo di 40 carte napoletane (4 semi; per ciascun seme 10 carte dall'asso al 7 più tre figure).

1. Qual'è la probabilità di estrarre un asso?

   Il numero di semi in un mazzo di carte napoletane è 4. Quindi ci sono 4 assi, di conseguenza la probabilità di trovare un asso è:

$$P(A) = \frac{4}{40} = \frac{1}{10} = 0.1 = 10\%$$

Estraiamo ora 2 carte.

2. Qual'è la probabilità di estrarre un asso e un re, rispettivamente?

   Calcolare le probabilità nel caso di:

   - Estrazione con reinserimento

     La probabilità di estrarre un asso e un re con reinserimento è data dal prodotto delle probabilità di estrarre un asso e un re:

$$P(A) = \frac{4}{40} = \frac{1}{10} = 0.1 = 10\% \quad \text{Estrarre un asso}$$

$$P(R) = \frac{4}{40} = \frac{1}{10} = 0.1 = 10\% \quad \text{Estrarre un re}$$

$$P(A \cap R) = P(A) \cdot P(R) = \frac{4}{40} \cdot \frac{4}{40} = \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{100} = 0.01 = 1\%$$

   - Estrazione senza reinserimento

     La probabilità di estrarre un re dopo aver estratto un asso è data dalla probabilità condizionata:

$$P(R \mid A) = \frac{P(R \cap A)}{P(A)} = \frac{\frac{1}{10} \cdot \frac{1}{10}}{\frac{1}{10}} = \frac{1}{10} = 0.1 = 10\%$$

1.4 Esercizio 4

Una popolazione si compone per un 40 percento di fumatori (F) e per il restante 60 per cento di non fumatori (N). Si sa che il 25 per cento dei fumatori e il 7 per cento dei non fumatori ha una malattia respiratoria cronica (M).

1. Calcolare la probabilità che un individuo scelto a caso sia effetto dalla malattia respiratoria.

   Un individuo scelto a caso ha una probabilità di essere fumatore malato:

   $$P(M \mid F) = 0.25$$

   e una probabilità di essere non fumatore malato:

   $$P(M \mid N) = 0.07$$

   La probabilità che un individuo scelto a caso sia affetto dalla malattia respiratoria è data dalla probabilità totale:

   $$P(M) = P(M \mid F) \cdot P(F) + P(M \mid N) \cdot P(N)$$

   $$P(M) = 0.25 \cdot 0.4 + 0.07 \cdot 0.6 = 0.1 + 0.042 = 0.142 = 14.2\%$$

2. Se l'individuo scelto è affetto dalla malattia, calcolare la probabilità che sia un fumatore.

   La probabilità che un individuo affetto dalla malattia sia un fumatore è data dalla probabilità condizionata:

   $$P(F \mid M) = \frac{P(M \mid F) \cdot P(F)}{P(M)} = \frac{0.25 \cdot 0.4}{0.142} = \frac{0.1}{0.142} = 0.704 = 70.4\%$$

1.5 Esercizio 5

Una particolare analisi del sangue è efficace al 99% nell'individuare una determinata malattia quando essa è presente. Si possono però verificare dei falsi positivi (ovvero una persona sana che si sottopone al test ha una probabilità pari a 0.01 di risultare erroneamente positiva al test). Se l'incidenza della malattia è del 0.5%, calcolare la probabilità che un individuo risultato positivo al test sia effettivamente malato.

$$T = \text{Test positivo}$$

$$M = \text{Persona malata} \quad \overline{M} = \text{Persona sana}$$

La probabilità che il test sia efficace è:

$$P(T \mid M) = 0.99$$

La probabilità che il test non sia efficace è:

$$P(T \mid \overline{M}) = 0.01$$

La probabilità di essere malato è:

$$P(M) = 0.005 = 0.5\%$$

La probabilità che il test sia positivo è data dalla probabilità totale:

$$P(T) = P(T \mid M) \cdot P(M) + P(T \mid \overline{M}) \cdot P(\overline{M})$$

$$P(T) = 0.99 \cdot 0.005 + 0.01 \cdot 0.995 = 0.00495 + 0.00995 = 0.0149 = 1.49\%$$

La probabilità che un individuo risultato positivo al test sia effettivamente malato è data dalla probabilità condizionata:

$$P(T \mid P) = \frac{P(T \mid M) \cdot P(M)}{P(T)}$$

$$P(T \mid P) = \frac{0.99 \cdot 0.005}{0.0149} = \frac{0.00495}{0.0149} \approx 0.332 = 33.2\%$$

1.6 Esercizio 6

In un vivaio si vendono dei sacchetti con 30 bulbi di dalie. I sacchetti sono di due tipi S_1 e S_2 . Tre quarti di tali sacchetti sono di tipo S_1 e contengono 10 bulbi di dalie rosse (R) e 20 bulbi di dalie gialle (G), mentre il restante quarto dei sacchetti è di tipo S_2 e contiene 5 bulbi di dalie rosse e 25 bulbi di dalie gialle.
Mario compra un sacchetto a caso e pianta un bulbo e vede di che colore è.

I sacchetti S_1 sono:

$$P(S_1) = \frac{3}{4} = 0.75 = 75\%$$

$$P(R \mid S_1) = \frac{10}{30} = \frac{1}{3} \approx 0.333 = 33.3\%$$

$$P(G \mid S_1) = \frac{20}{30} = \frac{2}{3} \approx 0.666 = 66.6\%$$

I sacchetti S_2 sono:

$$P(S_2) = \frac{1}{4} = 0.25 = 25\%$$

$$P(R \mid S_2) = \frac{5}{30} = \frac{1}{6} \approx 0.166 = 16.6\%$$

$$P(G \mid S_2) = \frac{25}{30} = \frac{5}{6} \approx 0.833 = 83.3\%$$

- Determinare la probabilità che il bulbo produca una dalia rossa.

  La probabilità che il bulbo produca una dalia rossa è data dalla probabilità totale:

  $$P(R) = P(R \mid S_1) \cdot P(S_1) + P(R \mid S_2) \cdot P(S_2)$$

  $$P(R) = \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{6} \cdot \frac{1}{4} = \frac{1}{4} + \frac{1}{24} = \frac{7}{24} \approx 0.291 = 29.1\%$$

- Se la dalia nata è rossa, qual è la probabilità che provenga da un sacchetto di tipo S_1?

  La probabilità che una dalia rossa provenga da S_1 è data dalla probabilità condizionata:
$$P(S_1 \mid R) = \frac{P(R \mid S_1) \cdot P(S_1)}{P(R)}$$
$$P(S_1 \mid R) = \frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{7}{24}} = \frac{1}{4} \cdot \frac{24}{7} = \frac{6}{7} \approx 0.857 = 85.7\%$$

1.7 Esercizio 7

In un test clinico, un individuo viene sottoposto ad un esame di laboratorio, per stabilire se ha o non ha una data malattia. Il test può avere esito positivo o negativo. C'è però sempre una possibilità di errore: può darsi che alcuni degli individui risultati positivi siano in realtà sani (falsi positivi), e che qualcuno degli individui risultati negativi siano in realtà malati (falsi negativi). Prima di applicare su larga scala un test nei laboratori, è quindi indispensabile valutarne la bontà, sottoponendo al test un campione di persone che sappiamo già se sono sane o malate. Uno dei parametri che definiscono la qualità diagnostica del test è la Sensibilità = 1 - P(falsi negativi).
In Italia c'è un malato di HIV ogni 40.000 persone. Un paziente si sottopone ad un test con una procedura che fornisce statisticamente lo 0,7% di falsi negativi e lo 0,01% di falsi positivi. Calcolare la probabilità a posteriori di essere ammalato, a test effettuato con esito positivo. Come cambia questa probabilità se però paziente e medico si convincono che, in base ai sintomi ed alle circostanze del possibile contagio, la probabilità a priori sia ad esempio 10 volte più alta della media nazionale?

$$M = \text{Persona malata} \quad \overline{M} = \text{Persona sana}$$
$$N = \text{Test negativo} \quad T = \text{Test positivo}$$

La probabilità di avere l'HIV in Italia è:
$$P(M) = \frac{1}{40000} = 0.000025 = 0.0025\%$$

La probabilità di non averlo è:
$$P(\overline{M}) = 1 - P(M) = 1 - 0.000025 = 0.999975 = 99.9975\%$$

La probabilità che il test sia falso negativo è:
$$P(N \mid M) = 0.007 = 0.7\%$$

La probabilità che il test sia falso positivo è:
$$P(T \mid \overline{M}) = 0.0001 = 0.01\%$$

La probabilità a posteriori di essere ammalato a test effettuato con esito positivo è data dalla probabilità condizionata:

$$P(M \mid T) = \frac{P(T \mid M) \cdot P(M)}{P(T)}$$

La probabilità che il test risulti positivo è data dalla probabilità totale:

$$P(T) = P(T \mid M) \cdot P(M) + P(T \mid \overline{M}) \cdot P(\overline{M})$$

$$P(T \mid M) = 1 - P(N \mid M) = 1 - 0.007 = 0.993 \quad \text{Sensibilità}$$

$$P(T) = 0.993 \cdot 0.000025 + 0.0001 \cdot 0.999975 \approx 0.000124822 = 0.0125\%$$

Se la probabilità a priori è 10 volte più alta della media nazionale:

$$P(M) = 10 \cdot \frac{1}{40000} = 0.00025 = 0.025\%$$

Le nuove probabilità che il test sia positivo è:

$$P(T) = 0.993 \cdot 0.00025 + 0.0001 \cdot 0.99975 \approx 0.000249822 = 0.0249\%$$

1.8 Esercizio 8

Si considerano due dadi apparentemente uguali, di cui uno equo (ossia fornisce un numero a caso con equiprobabilità) mentre l'altro fornisce un numero pari con probabilità doppia rispetto a quella dei numeri dispari.
Viene scelto a caso un dado e lo si lancia.

1. Calcolare la probabilità che esca un numero pari

$$Par = \text{Esce un numero pari}$$

$$Dis = \text{Esce un numero dispari}$$

$$P(D_1) = \frac{1}{2} = 0.5 = 50\% \quad \text{Dado equo}$$

$$P(D_2) = \frac{1}{2} = 0.5 = 50\% \quad \text{Dado non equo}$$

$$P(Par \mid D_1) = \frac{1}{2} = 0.5 = 50\%$$

$$P(Dis \mid D_1) = \frac{1}{2} = 0.5 = 50\%$$

$$P(Par \mid D_2) = \frac{6}{9} = \frac{2}{3} \approx 0.666 = 66.6\%$$

La probabilità che esca un numero pari è data dalla probabilità totale:

$$P(Par) = P(Par \mid D_1) \cdot P(D_1) + P(Par \mid D_2) \cdot P(D_2)$$

$$P(Par) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{4} + \frac{1}{3} = \frac{3}{12} + \frac{4}{12} = \frac{7}{12} \approx 0.583 = 58.3\%$$

2. Sapendo che è uscito un numero dispari, quanto vale la probabilità che il dado lanciato sia quello equo?

   La probabilità che il dado lanciato sia quello equo è data dalla probabilità condizionata:

$$P(D_1 \mid Dis) = \frac{P(Dis \mid D_1) \cdot P(D_1)}{P(Dis)}$$

$$P(D_1 \mid Dis) = \frac{\frac{1}{2} * \frac{1}{2}}{1 - 0.583} = \frac{\frac{1}{4}}{0.417} \approx 0.6 = 60\%$$

1.9 Esercizio 9

Consideriamo una popolazione di $N = 400$ piantine in cui si presentano i seguenti genotipi AA, Aa, aa. Supponiamo che una piantina venga estratta con uguale probabilità. Supponiamo inoltre che vi sono 196 piantine del genotipo AA, 168 del genotipo Aa, e 36 del genotipo aa.

$$N = 400$$

$$AA = 196$$

$$Aa = 168$$

$$aa = 36$$

1. Qual è la distribuzione dei genotipi, cioè se estraggo a caso una piantina qual è la probabilità di ciascun genotipo?

$$P(AA) = \frac{196}{400} = 0.49 = 49\%$$

$$P(Aa) = \frac{168}{400} = 0.42 = 42\%$$

$$P(aa) = \frac{36}{400} = 0.09 = 9\%$$

Supponiamo che l'allele a sia un letale recessivo, che però si esprime solo nella pianta allo stadio adulto.

2. La distribuzione precedente è ancora valida per la fase adulta (sì, no e perchè)?

   La distribuzione precedente non è valida per la fase adulta. Infatti, le piante con genotipo aa non sopravvivono allo stadio adulto, quindi la probabilità di trovare una pianta adulta con genotipo aa è nulla. Di conseguenza le probabilità in età adulta sono:

$$N = 400 - 36 = 364$$

$$P(AA) = \frac{196}{364} \approx 0.538 = 53.8\%$$

$$P(Aa) = \frac{168}{364} \approx 0.462 = 46.2\%$$

$$P(aa) = 0$$

3. Qual è la probabilità di trovare una pianta in vita dopo un certo periodo di tempo, quindi adulta?

   La probabilità di trovare una pianta in vita dopo un certo periodo di tempo è data dalla somma delle probabilità di trovare una pianta adulta con genotipo AA e Aa:

   $$P(AA) + P(Aa) = 0.49 + 0.42 = 0.91 = 91\%$$

4. Qual è la probabilità di estrarre una pianta adulta del genotipo AA? E del tipo Aa?

   $$P(AA) = 0.538 = 53.8\%$$
   $$P(Aa) = 0.462 = 46.2\%$$

1.10 Esercizio 10

Consideriamo la tabella che riporta i risultati di uno studio in cui sono stati osservati il numero di querce e lecci infestati da un certo insetto

	Infestato	Non infestato	Totali
Quercia	40	30	70
Leccio	52	10	62
Totali	92	40	132

1. Calcolare la probabilità che un albero sia infestato.

   $$P(I) = \frac{92}{132} \approx 0.697 = 69.7\%$$

2. Calcolare la probabilità che una quercia sia infestata.

   $$P(Q_I) = \frac{40}{70} \approx 0.571 = 57.1\%$$

3. Calcolare la probabilità che un leccio sia infestato.

   $$P(L_I) = \frac{52}{62} \approx 0.839 = 83.9\%$$

4. Quale definizione di probabilità si è utilizzata nei punti precedenti?

   La definizione utilizzata nei punti precedenti è: La probabilità è calcolata come il rapporto tra i casi favorevoli e i casi possibili.

1.11 Esercizio 11

Tre scrivanie tra loro indistinguibili contengono ciascuna due cassetti. La prima contiene una moneta d'oro in ciascun cassetto. La seconda una monta d'argento in un cassetto ed una moneta d'oro nell'altro, la terza due monte d'argento. Si apre un cassetto e si trova una moneta d'oro. Qual è la probabilità che anche nell'altro cassetto della stessa scrivania ci sia una moneta d'oro?

Avendo trovato una moneta d'oro, le uniche possibilità sono che la moneta sia stata trovata nella scrivania 1 o nella 2. La probabilità che la seconda moneta sia d'oro riduce le possibilità alla scrivania 1, quindi:

$$O = \text{Moneta d'oro}$$

$$P(S_1) = \frac{1}{3} = 0.333 = 33.3\%$$

$$P(O \mid S_1) = 1$$

$$P(O) = \left(1 + \frac{1}{2} + 0\right) \cdot \frac{1}{3} = \frac{1}{2} = 0.5 = 50\%$$

La probabilità che la seconda moneta sia d'oro è data dalla probabilità condizionata:

$$P(S_1 \mid O) = \frac{P(O \mid S_1) \cdot P(S_1)}{P(O)}$$

$$P(S_1 \mid O) = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} = 0.666 = 66.6\%$$

1.12 Esercizio 12

Da un'indagine nelle scuole, risulta che la percentuale degli alunni che portano gli occhiali è il 10% nelle scuole elementari, il 25% nelle scuole medie e il 40% nelle superiori.

$$O = \text{Porta gli occhiali}$$

$$E = \text{Frequenta le elementari}$$

$$M = \text{Frequenta le medie}$$

$$S = \text{Frequenta le superiori}$$

$$P(O \mid E) = 0.1 = 10\%$$

$$P(O \mid M) = 0.25 = 25\%$$

$$P(O \mid S) = 0.4 = 40\%$$

1. Calcolare la probabilità che scegliendo a caso 3 studenti, uno per fascia, almeno uno porti gli occhiali.

   La probabilità che almeno uno porti gli occhiali è data dal complementare dell'evento che nessuno porti gli occhiali:

   $$P(O \geq 1) = 1 - P(O = 0)$$

   $$P(O \geq 1) = 1 - \left(P(\overline{O} \mid E) \cdot P(\overline{O} \mid M) \cdot P(\overline{O} \mid S)\right) =$$
   $$= 1 - (0.9 \cdot 0.75 \cdot 0.6) = 1 - 0.405 = 0.595 = 59.5\%$$

2. Scegliendo uno studente a caso fra tutti (e supponendo che la scelta di ogni fascia sia equiprobabile) calcolare la probabilità che questo abbia gli occhiali.

   La probabilità che uno studente scelto a caso è data dalla probabilità totale:

$$P(O) = P(O \mid E) \cdot P(E) + P(O \mid M) \cdot P(M) + P(O \mid S) \cdot P(S)$$

$$P(O) = 0.1 \cdot \frac{1}{3} + 0.25 \cdot \frac{1}{3} + 0.4 \cdot \frac{1}{3} = \frac{1}{4} = 0.25 = 25\%$$

3. Sapendo che lo studente porti gli occhiali, calcolare la probabilità che frequenti le elementari.

   La probabilità che uno studente frequenti le elementari sapendo che porta gli occhiali è data dalla probabilità condizionata:

$$P(E \mid O) = \frac{P(O \mid E) \cdot P(E)}{P(O)}$$

$$P(E \mid O) = \frac{0.1 \cdot \frac{1}{3}}{\frac{1}{4}} \approx 0.2667 = 26.67\%$$

1.13 Esercizio 13

È stata studiata la relazione tra forma fisica e malattie cardiovascolari in un gruppo di impiegati delle ferrovie di sesso maschile, con questi risultati:

Frequenza cardiaca sotto esercizio (battiti/minuto)	Percentuale dei lavoratori	Mortalità per malattie cardiovascolari in 20 anni
≤ 105	20%	9.2%
106-115	30%	8.7%
116-127	30%	11.6%
> 127	20%	13.2%

Supponiamo che un certo test sia positivo se la frequenza cardiaca sotto esercizio è maggiore di 127 battiti al minuto e negativo altrimenti.

1. Qual è la probabilità di mortalità per malattie cardiovascolari in 20 anni?

$$T = \text{Test positivo}$$

$$\overline{T} = \text{Test negativo}$$

$$M = \text{Mortalità per malattie cardiovascolari in 20 anni}$$

$$F_1 = \text{Frequenza cardiaca} \leq 105$$

$$F_2 = \text{Frequenza cardiaca} \in [106, 115]$$

$$F_3 = \text{Frequenza cardiaca} \in [116, 127]$$

$$F_4 = \text{Frequenza cardiaca} > 127$$

La probabilità di mortalità per malattie cardiovascolari in 20 anni è data dalla probabilità totale:

$$P(M) = P(M \mid F_1) \cdot P(F_1) + P(M \mid F_2) \cdot P(F_2) +$$

$$P(M \mid F_3) \cdot P(F_3) + P(M \mid F_4) \cdot P(F_4)$$

$$P(M) = 0.092 \cdot 0.2 + 0.087 \cdot 0.3 + 0.116 \cdot 0.3 + 0.132 \cdot 0.2 = 0.1056 = 10.56\%$$

2. Qual è la probabilità di avere avuto un test positivo tra gli uomini che sono morti nel periodo di 20 anni?

   La probabilità di aver avuto un test positivo è data dalla probabilità condizionata:

   $$P(F_4 \mid M) = \frac{P(M \mid F_4) \cdot P(F_4)}{P(M)}$$

   $$P(F_4 \mid M) = \frac{0.132 \cdot 0.2}{0.1056} \approx 0.25 = 25\%$$

3. Qual è la probabilità di avere avuto un test positivo tra gli uomini che sono sopravvissuti nel periodo di 20 anni?

   La probabilità di avere avuto un test positivo tra gli uomini che sono sopravvissuti è data dalla probabilità condizionata:

   $$P(F_4 \mid \overline{M}) = \frac{P(\overline{M} \mid F_4) \cdot P(F_4)}{P(\overline{M})}$$

   $$P(F_4 \mid \overline{M}) = \frac{(1 - 0.132) \cdot 0.2}{1 - 0.1056} \approx 0.1941 = 19.41\%$$

4. Qual è la probabilità di morte tra gli uomini con un test negativo?

   La probabilità di morte tra gli uomini con un test negativo è data dalla probabilità totale dei casi in cui il test è negativo:

   $$P(M \mid \overline{T}) = P(M \mid F_1) \cdot P(F_1) + P(M \mid F_2) \cdot P(F_2) + P(M \mid F_3) \cdot P(F_3)$$

   $$P(M \mid \overline{T}) = 0.092 \cdot 0.2 + 0.087 \cdot 0.3 + 0.116 \cdot 0.3 = 0.0793 = 7.93\%$$

2 Variabili aleatorie: Binomiale, Poisson, Normale, Uniforme

2.1 Esercizio 1

Una persona arriva all'ascensore dell'edificio un minuto prima dell'orario di inizio del suo lavoro. Il tempo di attesa dell'ascensore, misurato in minuti, è una variabile aleatoria che ha una distribuzione uniforme in $[0, 3]$ (minuti). Considerando trascurabile il tempo impiegato dall'ascensore:

1. Qual'è la probabilità che la persona non arrivi in ritardo al lavoro?

   La probabilità che la persona arrivi in ritardo è:

   $$P(X > 1) = \frac{[1, 3]}{[0, 3]} = \frac{2}{3}$$

   La probabilità che la persona non arrivi in ritardo è:

   $$P(X \leq 1) = 1 - P(X > 1) = 1 - \frac{2}{3} = \frac{1}{3}$$

2. La probabilità precedente viene considerata bassa. Con quanto anticipo la persona deve arrivare prima all'ascensore in modo tale che la probabilità di arrivare in ritardo sia al massimo 0.05?

   La probabilità che la persona arrivi in ritardo è:

   $$P(X > x) = 0.05$$

   $$P(X > x) = 1 - P(X \leq x) = 1 - \frac{x}{3} = 0.05$$

   $$x = (1 - 0.05) \cdot 3 = 2.85 \text{ minuti}$$

2.2 Esercizio 2

Nel gioco del tiro al bersaglio si vincono 10 punti se si colpisce il bersaglio entro $2cm$ dal centro, 5 punti se si colpisce tra $2cm$ e $3cm$ e 2 punti se si colpisce tra $3cm$ e $5cm$. Oltre $5cm$ non si vince nulla.
Si supponga che la distanza del punto in cui si colpisce ed il centro del bersaglio abbia distribuzione uniforme sull'intervallo $(0, 10)$.

$$10 \text{ punti} = (0, 2]$$

$$5 \text{ punti} = (2, 3]$$

$$2 \text{ punti} = (3, 5]$$

$$0 \text{ punti} = (5, 10)$$

1. Qual'è il numero atteso di punti per lancio?

   Il numero atteso di punti per lancio è la media dei punti pesati per la probabilità di ottenere quel punteggio:

   $$\mathbb{E} = 10 \cdot \frac{2}{10} + 5 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} = 2 + 0.5 + 0.4 = 2.9$$

2. Qual'è la probabilità di avere 10 punti in un lancio?

   La probabilità di avere 10 punti in un lancio è:

   $$P(10) = \frac{2}{10} = \frac{1}{5} = 0.2 = 20\%$$

3. Assumendo i lanci indipendenti, qual'è la probabilità di ottenere 40 punti in 4 lanci?

   La probabilità di ottenere 40 punti in 4 lanci è data dalla probabilità di avere 10 punti in 1 lancio ripetuta 4 volte:

   $$P(40) = P(10)^4 = 0.2^4 = 0.0016$$

2.3 Esercizio 3

Arturo partecipa ad un gioco in cui la probabilità di vittoria è $p = \frac{2}{3}$.

1. Calcolare il numero minimo n di partite diverse che Arturo deve giocare, se vuole avere una probabilità superiore a $\frac{9}{10}$ di vincere almeno una volta.

   La probabilità di perdere una partita è $1-p = \frac{1}{3}$. La probabilità di vincere almeno una partita è il complementare di una sconfitta ripetuta n volte:

   $$P(X \geq 1) = 1 - P(X = 0) = 1 - (1-p)^n = \frac{9}{10}$$

   $$(1-p)^n = \frac{1}{10}$$

   $$n = \log_{1-p} \frac{1}{10} = \log_{\frac{1}{3}} \frac{1}{10} = \log_3 10 = 2.096$$

   Approssimando per eccesso all'intero più vicino, siccome non si possono giocare frazioni di partite, il numero minimo di partite da giocare è 3.

   $$n = 3 \text{ partite}$$

2. Supponendo che giochi $n = 5$ partite diverse, determinare la probabilità che in queste partite Arturo abbia ottenuto esattamente 3 vittorie.

   La probabilità che Arturo abbia ottenuto esattamente 3 vittorie in 5 partite è data dalla distribuzione binomiale:

   $$P(X = 3) = \binom{5}{3} \cdot p^3 \cdot (1-p)^2 = 10 \cdot \left(\frac{2}{3}\right)^3 \cdot \left(\frac{1}{3}\right)^2 \approx 0.33 = 33\%$$

3. Supponendo che giochi $n = 180$ partite diverse, stimare la probabilità che in queste partite Arturo abbia ottenuto un numero di vittorie compreso tra 108 e 121, inclusi.

   La probabilità che Arturo abbia ottenuto un numero di vittorie compreso tra 108 e 121 in 180 partite è data dalla distribuzione binomiale:

   $$P(108 \leq X \leq 121) = \sum_{k=108}^{121} \binom{180}{k} \cdot p^k \cdot (1-p)^{180-k}$$

   Questa probabilità è stata calcolata attraverso il seguente codice in linguaggio R:

```r
sum(dbinom(108:121, 180, 2 / 3))
# 0.5651441
```

$$P(108 \leq X \leq 121) = 0.565 = 56.5\%$$

2.4 Esercizio 4

Un esperimento consiste nel ripetere 10 volte l'estrazione (con reinserimento) di una biglia da un urna contenente 50 biglie, 15 delle quali sono bianche. Sia Y_{10} la variabile aleatoria che descrive il numero totale di biglie bianche estratte.

$$\text{Biglie totali} = 50$$

$$\text{Biglie bianche} = 15$$

$$p = \frac{15}{50} = 0.3$$

1. Determinare la probabilità che il numero di biglie bianche estratte siano almeno 2.

   La probabilità che il numero di biglie bianche estratte sia almeno 2 è data dalla distribuzione binomiale:

$$P(Y_{10} \geq 2) = 1 - P(Y_{10} < 2) = 1 - (P(Y_{10} = 0) + P(Y_{10} = 1))$$

$$\begin{aligned} P(Y_{10} \geq 2) = 1 - \left(\binom{10}{0} \cdot (0.3)^0 \cdot (0.7)^{10} + \binom{10}{1} \cdot (0.3)^1 \cdot (0.7)^9\right) &= \\ = 1 - \left(\frac{10!}{10!} \cdot 1 \cdot (0.7)^{10} + \frac{10!}{9!} \cdot 0.3 \cdot (0.7)^9\right) &= \\ = 1 - \left((0.7)^{10} + 3 \cdot (0.7)^9\right) & \\ = 1 - (0.149308346) \approx 0.85 = 85\% & \end{aligned}$$

2. Come si calcola la probabilità che il numero di biglie estratte sia maggiore strettamente di 3 e minore o uguale di 6? ($P(3 < Y_{10} \leq 6)$)

   La probabilità si calcola sommando le probabilità di avere 4 biglie bianche, 5 biglie bianche e 6 biglie bianche:

$$P(3 < Y_{10} \leq 6) P(Y_{10} = 4) + P(Y_{10} = 5) + P(Y_{10} = 6)$$

$$\begin{aligned} P(3 < Y_{10} \leq 6) = \left(\binom{10}{4} \cdot 0.3^4 \cdot 0.7^6\right) &+ \\ \left(\binom{10}{5} \cdot 0.3^5 \cdot 0.7^5\right) &+ \\ \left(\binom{10}{6} \cdot 0.3^6 \cdot 0.7^4\right) &= \\ = 0.200120949 + 0.102919345 + 0.036756909 \approx 0.34 = 34\% & \end{aligned}$$

2.5 Esercizio 5

Un calcolatore è collegato ad una rete che permette l'accesso ad un massimo di 20 persone. Collegati a questa rete vi sono i terminali di 24 operatori, ognuno dei quali, ad un dato istante, richiede con probabilità $p = 0.6$ di essere connesso al calcolatore centrale.
Qual'è la probabilità che ad un dato istante la rete sia satura (cioè che tutti e 20 gli accessi siano utilizzati)?

La probabilità che la rete sia satura è data dalla distribuzione binomiale:

$$P(X \geq 20) = \sum_{k=20}^{24} \binom{24}{k} \cdot p^k \cdot (1-p)^{24-k}$$

$$\begin{aligned}
P(X = 20) &= \binom{24}{20} \cdot 0.6^{20} \cdot 0.4^4 = \frac{24!}{20! \cdot 4!} \cdot 0.000036562 \cdot 0.0256 = \\
&= 10626 \cdot 0.000036562 \cdot 0.0256 \approx 0.0099 \\
P(X = 21) &= \binom{24}{21} \cdot 0.6^{21} \cdot 0.4^3 = \frac{24!}{21! \cdot 3!} \cdot 0.000021937 \cdot 0.064 = \\
&= 2024 \cdot 0.000021937 \cdot 0.064 \approx 0.0028 \\
P(X = 22) &= \binom{24}{22} \cdot 0.6^{22} \cdot 0.4^2 = \frac{24!}{22! \cdot 2!} \cdot 0.000013162 \cdot 0.16 = \\
&= 276 \cdot 0.000013162 \cdot 0.16 \approx 0.00058 \\
P(X = 23) &= \binom{24}{23} \cdot 0.6^{23} \cdot 0.4^1 = \frac{24!}{23! \cdot 1!} \cdot 0.000007897 \cdot 0.4 = \\
&= 24 \cdot 0.000007897 \cdot 0.4 \approx 0.00008 \\
P(X = 24) &= \binom{24}{24} \cdot 0.6^{24} \cdot 0.4^0 = 0.6^{24} \approx 0.000004
\end{aligned}$$

$$P(X \geq 20) = 0.0099 + 0.0028 + 0.00058 + 0.00008 + 0.000004 \approx 0.0135 = 1.35\%$$

2.6 Esercizio 6

In una linea produttiva ogni pezzo ha probabilità $p = 0.03$ di essere difettoso. L'essere difettoso è indipendente da pezzo a pezzo.
Calcolare la probabilità che su 30 pezzi non più di 2 siano difettosi.

La probabilità che su 30 pezzi non più di 2 siano difettosi è data dalla distribuzione binomiale:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

Per semplicità di calcolo uso una funzione di R:

```r
pbinom(2, 30, 0.03)
# 0.9399309
```

La probabilità che su 30 pezzi non più di 2 siano difettosi è:

$$P(X \leq 2) = 0.9399 = 93.99\%$$

2.7 Esercizio 7

Una segreteria di un'azienda di statistica immette i dati in una banca dati con un ritmo medio di 200 immissioni all'ora. Si suppone indipendenza tra le immissioni.
Quale modello probabilistico si può utilizzare per calcolare la probabilità che ci siano esattamente 5 immissioni nel successivo minuto? Si può ipotizzare che sia una Poisson? Determinare tale probabilità.

Il modello probabilistico più adatto è la distribuzione di Poisson perchè abbiamo degli eventi indipendenti che si verificano con una certa frequenza in un certo intervallo.
La probabilità che ci siano esattamente 5 immissioni nel successivo minuto è data dalla distribuzione di Poisson:

$$\mathcal{P}_\lambda(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

$$\lambda = 200 \text{ immissioni/ora} = \frac{200}{60} \text{ immissioni/minuto} = 3.33 \text{ immissioni/minuto}$$

$$n = 5 \text{ immissioni}$$

$$\mathcal{P}(5) = \frac{3.33^5}{5!} e^{-3.33} \approx 0.122$$

2.8 Esercizio 8

Un esperimento consiste nel ripetere 60 volte l'estrazione (con reinserimento) di una biglia da un'urna contenente 50 biglie, 15 delle quali sono bianche. Sia Y_{60} la variabile aleatoria che descrive il numero totale di biglie bianche estratte. Determinare la probabilità che il numero di bilglie estratte sia maggiore strettamente di 10 e minore o uguale di 30. ($P(10 < Y_{60} \leq 30)$)

La probabilità che il numero di biglie estratte sia maggiore strettamente di 10 e minore o uguale di 30 è data dalla distribuzione binomiale:

$$P(10 < Y_{60} \leq 30) = \sum_{k=11}^{30} \binom{60}{k} \cdot p^k \cdot (1-p)^{60-k}$$

$$P(10 < Y_{60} \leq 30) = \sum_{k=11}^{30} \binom{60}{k} \cdot 0.3^k \cdot 0.7^{60-k}$$

```
sum(dbinom(11:30, 60, 15 / 50))
# 0.9857581
```

$$P(10 < Y_{60} \leq 30) \approx 0.9857 = 98.57\%$$

2.9 Esercizio 9

In un corso di laurea a numero chiuso si ritiene che il numero di studenti non debba essere superiore a 80. Inoltre è noto, però, che il 25% degli studenti che superano l'esame di ammissione cambiano idea e non confermano l'iscrizione.

Se 100 studenti superano l'esame di ammissione, qualè la probabilità che più di 80 decidano di iscriversi?

La probabilità che uno studente che ha superato l'esame si iscriva è:

$$P(I) = 0.75$$

La probabilità che più di 80 studenti si iscrivano è data dalla distribuzione binomiale:

$$P(X > 80) = \sum_{k=81}^{100} \binom{100}{k} \cdot 0.75^k \cdot 0.25^{100-k}$$

```
pbinom(81, 100, 0.75, lower.tail=FALSE)
# 0.06301142
```

$$P(X > 80) \approx 0.063 = 6.3\%$$

2.10 Esercizio 10

Da una rilevazione risulta che il numero di incidenti sul lavoro avvenuti in un'azienda in un mese è una variabile distribuita secondo Poisson con valor medio $\lambda = 1.5$. Calcolare:

$$\mathcal{P}_\lambda(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

1. La probabilità che in un mese non ci siano incidenti.

   La probabilità che in un mese non ci siano incidenti è data dalla distribuzione di Poisson:

   $$\mathcal{P}(0) = \frac{1.5^0}{0!} e^{-1.5} = e^{-1.5} \approx 0.2231 = 22.31\%$$

2. La probabilità che in un mese ci siano più di 2 incidenti.

   La probabilità che in un mese ci siano più di 2 incidenti è data dal complementare della probabilità che in un mese ci siano al massimo 2 incidenti:

   $$P(X > 2) = 1 - P(X \leq 2) = 1 - (P(X = 0) + P(X = 1) + P(X = 2))$$

   $$P(X > 2) = 1 - \left(0.2231 + \frac{1.5^1}{1!} e^{-1.5} + \frac{1.5^2}{2!} e^{-1.5}\right) =$$

   $$= 1 - \left(0.2231 + 1.5e^{-1.5} + 1.125e^{-1.5}\right) \approx 0.1912 = 19.12\%$$

2.11 Esercizio 11

Da rilevazioni statistiche si è valutato che una persona su 100.000 è allergica ad un tipo di polline. In una popolazione di 300.000 abitanti calcola la probabilità che:

1. Nessuna persona sia allergica.

   Questo problema può essere risolto con la distribuzione di Poisson:

   $$\lambda = 300000 \cdot \frac{1}{100000} = 3$$

   La probabilità che nessuno sia allergico è:

   $$\mathcal{P}(0) = \frac{3^0 \cdot e^{-3}}{0!} = e^{-3} \approx 0.0498 = 4.98\%$$

2. Al massimo 3 persone siano allergiche.

   La probabilità che al massimo 3 persone siano allergiche è data dalla distribuzione di Poisson:

   $$\mathcal{P}(X \leq 3) = \sum_{k=0}^{3} \frac{3^k \cdot e^{-3}}{k!}$$

   Il risultato della sommatoria è dato dal seguente codice in R:

   ```
   ppois(3, 3)
   # 0.6472319
   ```

   $$\mathcal{P}(X \leq 3) \approx 0.6472 = 64.72\%$$

2.12 Esercizio 12

La probabilità che un individuo contragga una malattia rara è dello 0.3 per mille individui.
Qual è la probabilità che in una città dove vivono 20000 persone vi siano meno di 4 persone che contraggono questa malattia?

$$p = \frac{0.3}{1000} = 0.0003$$

Questo problema è modellabile con la distribuzione di Poisson:

$$\lambda = 20000 \cdot 0.0003 = 6$$

La probabilità che vi siano meno di 4 persone che contraggono la malattia è:

$$\mathcal{P}(X < 4) = \sum_{k=0}^{3} \frac{6^k \cdot e^{-6}}{k!}$$

Il risultato della sommatoria è dato dal seguente codice in R:

```r
ppois(3, 6)
# 0.1512039
```

$$\mathcal{P}(X < 4) \approx 0.1512 = 15.12\%$$

2.13 Esercizio 13

Alla focacceria arrivano mediamente 30 clienti l'ora. Calcolare la probabilità che in due minuti:

1. Non arrivi nessuno.

   Questo problema è modellabile con la distribuzione di Poisson:

   $$\lambda = 30 \cdot \frac{2}{60} = 1$$

   La probabilità che non arrivi nessuno è data da:

   $$\mathcal{P}(0) = \frac{1^0 \cdot e^{-1}}{0!} = e^{-1} \approx 0.3679 = 36.79\%$$

2. Arrivino almeno 3 clienti.

   La probabilità che arrivino almeno 3 clienti è data dal complementare della probabilità che arrivino al massimo 2 clienti:

   $$\mathcal{P}(X \geq 3) = 1 - \mathcal{P}(X < 3) = 1 - \sum_{k=0}^{2} \frac{1^k \cdot e^{-1}}{k!}$$

   Il risultato della sommatoria è dato dal seguente codice in R:

   ```r
   1 - ppois(2, 1)
   # 0.0803014
   ```

   $$\mathcal{P}(X \geq 3) \approx 0.0803 = 8.03\%$$

2.14 Esercizio 14

Ad un esame universitario, il voto X medio è stato $\mu = \mathbb{E}[X] = 24$, con $\sigma^2 = var(X) = 4^2$. Supponendo i voti normalmente distribuiti, calcolare la probabilità che uno studente abbia riportato:

1. Un voto superiore a 27.

   Questo problema è modellabile con la distribuzione normale standard, bisogna quindi standardizzare il voto:

   $$Z = \frac{X - \mu}{\sigma} = \frac{27 - 24}{4} = 0.75$$

Per trovare la probbailità di trovare un voto superiore a 27 bisogna cercare nelle tabelle dei valori della distribuzione normale standard e prendere il complementare:

$$P(Z > 0.75) = 1 - 0.7734 = 0.2266 = 22.66\%$$

2. Un voto non inferiore a 22.

   Per trovare la probabilità di trovare un voto non inferiore a 22 bisogna standardizzare il voto:

   $$Z = \frac{X - \mu}{\sigma} = \frac{22 - 24}{4} = -0.5$$

   Successivamente bisogna cercare il valore di probabilità nella tabella e prendere il complementare:

   $$P(Z \geq -0.5) = 1 - 0.3085 = 0.6915 = 69.15\%$$

2.15 Esercizio 15

L'altezza X di un gruppo di 20.000 individui è distribuita normalmente con media $\mu = \mathbb{E}[X] = 170cm$ e con deviazione standard $\sigma = \sqrt{var(X)} = 10cm$.

1. Qual è la probabilità che l'altezza sia compresa fra $155cm$ e $180cm$?

   Questo problema è modellabile con la distribuzione normale standard, bisogna quindi standardizzare le altezze:

   $$Z_1 = \frac{155 - 170}{10} = -1.5$$
   $$Z_2 = \frac{180 - 170}{10} = 1$$

   La probabilità che l'altezza sia compresa fra $155cm$ e $180cm$ è data dalla differenza delle probabilità trovate dalle tabelle della distribuzione normale standard:

   $$P(-1.5 < Z < 1) = P(Z < 1) - P(Z < -1.5)$$
   $$P(-1.5 < Z < 1) = 0.8413 - 0.0668 = 0.7745 = 77.45\%$$

2. Quante persone ci si aspetta che siano alte almeno 2 metri?

   Standardizzando l'altezza:

   $$Z = \frac{200 - 170}{10} = 3$$

   La probabilità che una persona sia alta almeno 2 metri è:

   $$P(Z > 3) = 1 - 0.9987 = 0.0013 = 0.13\%$$

   Per sapere quante persone ci si aspetta che siano alte almeno 2 metri bisogna moltiplicare la probabilità per il numero di persone:

   $$0.0013 \cdot 20000 = 26$$

   Ci si aspetta, quindi, che 26 persone siano alte almeno 2 metri in un gruppo di 20.000 individui

3. Quante persone ci si aspetta che siano alte non più $160cm$?

   Standardizzando l'altezza:

   $$Z = \frac{160 - 170}{10} = -1$$

   La probabilità che una persona sia alta non più di $160cm$ è:

   $$P(Z < -1) = 0.1587$$

   Il numero di persone che ci si aspetta siano alte non più di $160cm$ è:

   $$0.1587 \cdot 20000 = 3174$$