

Probabilità e Statistica

Esercizi

UniVR - Dipartimento di Informatica

Fabio Irimie

2° Semestre 2023/2024

Indice

1	Probabilità elementari e probabilità condizionate	2
1.1	Esercizio 1	2
1.2	Esercizio 2	2
1.3	Esercizio 3	3
1.4	Esercizio 4	4
1.5	Esercizio 5	4
1.6	Esercizio 6	5
1.7	Esercizio 7	6
1.8	Esercizio 8	7
1.9	Esercizio 9	8
1.10	Esercizio 10	9
1.11	Esercizio 11	9
1.12	Esercizio 12	10
1.13	Esercizio 13	11

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 $10!$:

$$\text{card}(\Omega) = 10!$$

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

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

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:

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

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

1.3 Esercizio 3

Consideriamo l'estrazione di una carta da un 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:

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

1.4 Esercizio 4

Una popolazione si compone per un 40 per cento 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 affetto dalla malattia respiratoria.

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

$$P(M | F) = 0.25$$

e una probabilità di essere non fumatore malato:

$$P(M | N) = 0.07$$

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

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

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

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:

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

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 = Test positivo

M = Persona malata \overline{M} = Persona sana

La probabilità che il test sia efficace è:

$$P(T | M) = 0.99$$

La probabilità che il test non sia efficace è:

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

La probabilità di essere malato è:

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

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

$$P(T) = P(T|M) \cdot P(M) + P(T|\bar{M}) \cdot P(\bar{M})$$

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

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

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

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

1.6 Esercizio 6

In un vivaio si vendono dei sacchetti con 30 bulbi di dalia. 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 dalia rosse (R) e 20 bulbi di dalia gialle (G), mentre il restante quarto dei sacchetti è di tipo S_2 e contiene 5 bulbi di dalia rosse e 25 bulbi di dalia gialle.

Mario compra un sacchetto a caso e pianta un bulbo e vede di che colore è.

I sacchetti S_1 sono:

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

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

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

I sacchetti S_2 sono:

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

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

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

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

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

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

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

- 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:

$$P(S_1 | R) = \frac{P(R | S_1) \cdot P(S_1)}{P(R)}$$

$$P(S_1 | R) = \frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{7}{24}} = \frac{1}{4} \cdot \frac{24}{7} = \frac{6}{7} \approx 0.857 = 85.7\%$$

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 = $1 - P(\text{falsi negativi})$.

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:

$$P(N | M) = 0.007 = 0.7\%$$

La probabilità che il test sia falso positivo è:

$$P(T | \overline{M}) = 0.0001 = 0.01\%$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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:

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

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

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:

$$N = 400 - 36 = 364$$

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

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

$$P(aa) = 0$$

- 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:

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

- 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

- Calcolare la probabilità che un albero sia infestato.

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

- Calcolare la probabilità che una quercia sia infestata.

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

- Calcolare la probabilità che un leccio sia infestato.

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

- 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 moneta d'argento in un cassetto ed una moneta d'oro nell'altro, la terza due monete 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:

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

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

$$P(O | S_1) = 1$$

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

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

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

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

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.

$O = \text{Porta gli occhiali}$

$E = \text{Frequenta le elementari}$

$M = \text{Frequenta le medie}$

$S = \text{Frequenta le superiori}$

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

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

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

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:

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

$$\begin{aligned} P(O \geq 1) &= 1 - (P(\bar{O} | E) \cdot P(\bar{O} | M) \cdot P(\bar{O} | S)) = \\ &= 1 - (0.9 \cdot 0.75 \cdot 0.6) = 1 - 0.405 = 0.595 = 59.5\% \end{aligned}$$

- 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:

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

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

- 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:

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

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

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.

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

T = Test positivo

\bar{T} = Test negativo

M = Mortalità per malattie cardiovascolari in 20 anni

F_1 = Frequenza cardiaca ≤ 105

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

F_3 = 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:

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

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

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

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:

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

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

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:

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

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

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:

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

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