Soluzioni di esercizi del libro

a cura di Giovanni M. Marchetti

2016 ver. 1.0

Indice

Capitolo 1. Introduzione

Esercizio 1.9

a. Grafici e misure di sintesi.
b. Per esempio, la stima della soddisfazione media.
c. Per esempio, la relazione tra media e intenzione ad iscriversi.

Capitolo 2. Descrizione grafica dei dati

Esercizio 2.1

a. Variabile numerica che deriva da un processo di conteggio: quantitativa discreta.
b. Variabile numerica espressa in valuta: quantitativa continua.
c. Come per b.
d. Variabile categorica, le cui categorie hanno una gerarchia tra le modalità: qualitativa ordinale.

Esercizio 2.5

a. Variabile categorica, qualitativa nominale
b. Variabile numerica, quantitativa discreta
c. Variabile categorica, qualitativa nominale
d. Variabile categorica, qualitativa nominale

Esercizio 2.7

a. Hours each time: è una variabile numerica ricodificata in classi, perciò rappresenta una variabile qualitativa ordinale.
b. Tutte le variabili E-H. NB: le variabili dicotomiche o dummy (che possono assumere solo due modalità) si considerano sempre qualitative nominali.
c. Days/Week: il numero di giorni a settimana rappresenta una variabile quantitativa discreta.

Esercizio 2.13

a. Si tratta di una variabile quantitativa discreta. Per la distribuzione di frequenze si consiglia di costruire una tabella, con il carattere ricodificato in classi, indicando nella prima colonna le classi (in ordine crescente) e nella seconda colonna la frequenza di ogni modalità.

Occorre innanzitutto scegliere il numero delle classi, tra 5 e 7. Si scelga per esempio un numero di classi pari a 5 (ma è corretto anche scegliere un numero diverso di classi!).

Per determinare l'ampiezza di ciascuna classe, occorre applicare la formula:

$$w = \frac{\text{Massimo} - \text{Minimo}}{\text{N.Classi}} = (65 - 12)/5 = 53/5 = 10.6$$

È importante arrotondare sempre per eccesso l'ampiezza determinata con la formula (in questo caso 11).

Le classi saranno dunque:

`[12;23) [23;34) [34;45) [45;56) [56;67)`

La distribuzione di frequenze richiesta è:

Classi	Frequenze	F. Relative%
$12 \vdash 23$	6	21
$23 \vdash 34$	3	11
$34 \vdash 45$	10	36
$45 \vdash 56$	2	7
$56 \vdash 67$	7	25
	28	100

Capitolo 3. Descrizione numerica dei dati

Esercizio 3.1

$$\bar{x} = \frac{20 + 73 + 75 + 80 + 82}{5} = 66$$

Posizione centrale nella successione ordinata = (5+1)/3 = 3. Mediana = 75.

La moda non riusciamo a calcolarla perché tutte le modalità hanno frequenza uguale.

Delle tre misure di sintesi, quella che descrive meglio i dati è la mediana, poiché esiste un dato anomalo che sposta la media eccessivamente verso il basso.

Esercizio 3.5

La media è $\bar{x} = 20.74$. Per la mediana prima ordino i dati

`10.2 13.1 15.0 15.8 16.9 17.3 18.2 24.7 25.3 28.4 29.3 34.7`

Quindi trovo la posizione centrale in questa graduatoria. Poiché il numero di unità è pari, ci sono due posizioni centrali. Le trovo calcolando $(n + 1)/2 = (12 + 1)/2 = 6.5$. Quindi le posizioni sono la 6 e la 7. La mediana è la semisomma del sesto e del settimo valore in graduatoria. Quindi Mediana $= (17.3 + 18.2)/2 = 17.75$.

Esercizio 3.14

Il coefficiente di variazione è il rapporto tra deviazione standard e il valore assoluto della media. Occorre dunque anzitutto calcolare queste misure di sintesi. La media è $\bar{x} = 9$. La deviazione standard è (sono dati campionari e quindi si usa il denominatore $(n-1)$)

$$s = \sqrt{\frac{10}{4}} = 1.5811.$$

Perciò il coefficiente di variazione è

$$\frac{s}{|\bar{x}|} = 1.5811/9 = 0.1757 = 17.57\%.$$

Esercizio 3.15

Il numero di osservazioni è $n = 24$. La successione in ordine crescente è

`12, 13, 13, 13, 14, 16, 20, 20, 23, 25, 26, 27, 27, 28, 29, 35, 37, 37, 40, 40, 40, 45, 49, 66`

a. La media è $\bar{x} = 28.9583$ secondi.

b. La deviazione standard è $s = 13.5694$ secondi.

c. Il minimo e il massimo sono 12 e 66. La mediana è la semisomma delle osservazioni al 12mo e 13mo posto. Siccome sono entrambi uguali a 27 la mediana è 27. I quartili si possono calcolare analogamente con una formula spiegata sul libro. Siccome non è in programma non mi soffermo sui dettagli.

d. Il coefficiente di variazione (non confondete con il campo di variazione!) è

$$CV = \frac{s}{|\bar{x}|} = 13.5694/28.9583 = 46.86\%.$$

Esercizio 3.27

I dati rono raggruppati in classi. Quindi si calcolano i valori approssimati di media e varianza definendo i punti centrali di classe c_i e risulta $\bar{x} = \sum_i c_i n_i / n = 505/40$. La tabella seguente è utile per fare i calcoli della varianza

Classi	Freq	c_i	Freq $\times$ c_i	$c_i - \bar{x}$	$(c_i - \bar{x})^2$	$(c_i - \bar{x})^2 n_i$
$0-4$	5	2	10	-10.625	112.891	564.4531
$5-9$	8	7	56	-5.625	31.641	253.1250
$10-14$	11	12	132	-0.625	0.391	4.2969
$15-19$	9	17	153	4.375	19.141	172.2656
$20-24$	7	22	154	9.375	87.891	615.2344
	40		505			1609.3750

La varianza si calcola come

$$\sum_i (c_i - \bar{x})^2 n_i / n = 1609.375/39 = 41.2661$$

dove abbiamo usato il denominatore $n-1$. La deviazione standard è 6.42.

Esercizio 3.30

a. Somma del numero di richieste x, numero di polizze $= 13 \cdot 1 + 2 \cdot 5 + 3 \cdot 4 + 4 \cdot 2 + 5 \cdot 3 + 6 \cdot 2 = 70$. Quindi

$$\bar{x} = \frac{70}{50} = 1.4 \text{ polizze.}$$

Per calcolare la varianza è meglio usare

$$\sigma^2 = \sum x_i^2 n_i / N - \bar{x}^2$$

Quindi

$$\sigma^2 = \frac{1 \times 13 + 4 \times 5 + 9 \times 4 + 16 \times 2 + 25 \times 3 + 36 \times 2}{50} - 1.4^2$$

e quindi $\sigma^2 = \frac{258}{50} - 1.4^2 = 3$.

Poiché i dati sono campionari si calcola la varianza corretta

$$s^2 = \sigma^2 \frac{N}{N-1}.$$

Quindi in questo caso è $s^2 = 3\frac{50}{49} = 3.061$ e la deviazione standard è $s = \sqrt{3.061} = 1.75$.

Esercizio 3.38

```
x:  5, 21, 14, 11,  9,  4,  7, 21, 17, 14,  9,  7,  9, 21, 13, 14,  9, 4
y: 53, 65, 48, 66, 46, 56, 53, 57, 49, 66, 54, 56, 53, 52, 49, 56, 59, 56
```

Quindi

$\bar{x} = 11.61, \bar{y} = 55.22$.

Il prodotto degli scarti $(x_i - \bar{x})(y_i - \bar{y})$ è

```
14.69, 91.80, -17.25, -6.59, 24.08, -5.92, 10.25, 16.69, -33.53, 25.75, 3.19, -3.59, 5.80, -30.25, -8.64, 1.8
```

La covarianza è (con denominatore N)

$$\sum_i [(x_i - \bar{x})(y_i - \bar{y})]/18 = 72.55/18 = 4.03.$$

Varianza di X: Quadrati degli scarti $(x_i - \bar{x})^2$:

```
43.71, 88.15, 5.71, 0.37, 6.82, 57.93, 21.26, 88.15, 29.04, 5.71, 6.82, 21.26, 6.82, 88.15, 1.93, 5.71, 6.82,
```

Quindi

$$\sum_i (x_i - \bar{x})^2 = 544.28$$

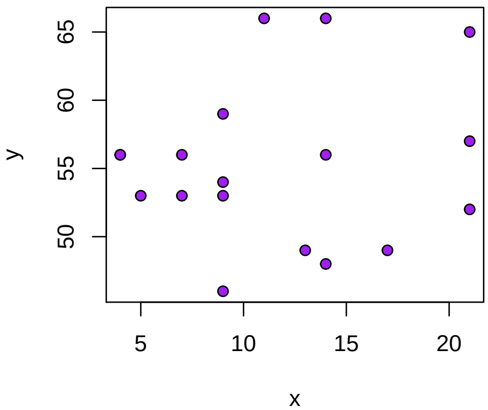
e la varianza di y è $542.28/18 = 30.12$ (sempre con denominatore N).

Analogamente si ottiene la $\sigma_Y^2 = 32.73$.

Quindi il coefficiente di correlazione è

$$\rho_{xy} = 4.03/\sqrt{30.12 \times 32.73} = 0.128.$$

c. L'analisi mostra che le due variabili sono quasi incorrelate. Guardate lo scatter:



L'associazione tra la dose del farmaco e la guarigione è debole. Quindi supponendo che i farmaco sia sempre potenzialmente dannoso si potrebbe pensare di prescrivere un dosaggio zero!

Capitolo 4. Probabilità

Esercizio 4.5

f. L'intersezione $A \cap B$ è l'evento che ci vogliano 5 giorni. L'evento $\bar{A} \cap B$ è l'evento che ci vogliano 4,3,2,1 giorni.

Quindi l'unione dei due è l'evento che ci vogliano 1,2,3,4,5 giorni cioè < 6 giorni. Questa è la esattamente la definizione dell'evento B.

g. $\bar{A} \cap B$ è l'evento che ci vogliano 4,3,2,1 giorni. A è l'evento che ci vogliano 5,6,7 giorni. Quindi l'unione dei due è l'evento 1,2,3,4,5,6,7 giorni. Questo è l'evento $A \cup B$. Quindi $A \cup (\bar{A} \cap B) = A \cup B$. Nota che per la proprietà distributiva

$$A \cup (\bar{A} \cap B) = (A \cup \bar{A}) \cap (A \cup B) = S \cap (A \cup B) = A \cup B.$$

Quindi non c'è bisogno di tutti questi conti.

Esercizio 4.7

a. Lo spazio campionario è il seguente

```
1. M1M2  5. M2M1   9. M3M1  13. T1M1  17. T2M1
2. M1M3  6. M2M3  10. M3M2  14. T1M2  18. T2M2
3. M1T1  7. M2T1  11. M3T1  15. T1M3  19. T2M3
4. M1T2  8. M2T2  12. M3T2  16. T1T2  20. T2T1
```

Ci sono cioè 20 eventi elementari che abbiamo numerato. Non sono 25 perché non possono capitare T1T1, T2T2, ecc.

b. Allora

$$A = \{3, 4, 7, 8, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20\}$$

c.

$$B = \{1, 2, 5, 6, 9, 10, 16, 20\}$$

d.

$$\bar{A} = \{1, 2, 5, 6, 9, 10\}$$

e. $A \cap B = \{16, 20\}$ è l'evento che si verifica se viene scelta almeno una Toyota e le due auto hanno la stessa marca.

Invece $\bar{A} \cap B = \{1, 2, 5, 6, 9, 10\}$. Quindi la loro unione è B.

f. Abbiamo che l'unione di A e $\bar{A} \cap B$ è

$$\{3, 4, 7, 8, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 1, 2, 5, 6, 9, 10\}$$

cioè eliminando i doppioni

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20\}$$

ossia tutti gli eventi elementari. Questa coincide con $A \cup B$.

NOIOSISSIMO!

Esercizio 4.8

Lo spazio campionario è `AAAAABBBBBBB`. L'evento che un sottoinsieme di 2 lettere contenga una A e una B è $\{A, B\}$. I casi possibili sono

$$\binom{5}{1}\binom{7}{1} = 35.$$

Cioè posso scegliere la A in 5 modi possibili e per ciascuna posso scegliere la B in 7 modi possibili. Ricordate l'esercizio in cui bisognava pescare a caso un comitato di maschi e femmine. Lo spazio campionario contiene

$$\binom{12}{2} = (12)(11)/2 = 66$$

casi possibili. Quindi la probablità è $35/66 = 0.53$.

Per visualizzare il tutto potete costruire gli eventi elementari con una tabella:

```
  | A  A  A  A  A  B  B  B  B  B  B  B
--+-----------------------------------
A |AA AA AA AA AA AB AB AB AB AB AB AB
A |   AA AA AA AA AB AB AB AB AB AB AB
A |      AA AA AA AB AB AB AB AB AB AB
A |         AA AA AB AB AB AB AB AB AB
A |            AA AB AB AB AB AB AB AB
B |               BB BB BB BB BB BB BB
B |                  BB BB BB BB BB BB
B |                     BB BB BB BB BB
B |                        BB BB BB BB
B |                           BB BB BB
B |                              BB BB
B |                                 BB
```

Da dove vedete che gli AB sono 35 e tutti i casi possibili sono 66.

Esercizio 4.9

È più difficile del precedente, ma si usa la stessa regola. Lo spazio campionario è `AAAAAABBBB`.

I casi possibili sono

$$\binom{10}{3} = (10 \cdot 9 \cdot 8)/(3 \cdot 2) = 120.$$

I casi favorevoli sono (usando la regola di prima)

$$\binom{6}{1}\binom{4}{2} = 6 \cdot (4 \cdot 3)/2 = 36.$$

Quindi la probabilità è $36/120 = 0.3$.

Esercizio 4.10

Stesso sistema: l'insieme dei casi possibili è $\binom{16}{4} = 1820$. Casi favorevoli: $\binom{10}{2}\binom{6}{2} = 675$. Quindi la probabilità è $675/1820 = 0.37$.

Esercizio 4.11

Questo è facile: la probabilità è uguale alla frequenza relativa di norvegesi $= 20000/120000 = 1/6$.

Esercizio 4.12

Si usa regola $P(A \cap B) = P(A)P(B \mid A)$ dove $A =$ norvegese alla prima estrazione e $B =$ norvegese alla seconda estrazione:

$$\begin{aligned} P(\text{2 norvegesi su 2}) &= P(\text{norvegese alla prima }) \cdot P(\text{norvegese alla seconda } \mid \text{norvegese alla prima}) \\ &= (20000/180000)(19999/179999) = 0.01234. \end{aligned}$$

Nota che alla seconda estrazione la popolazione è praticamente uguale.

Esercizio 4.14

a. $P(A) = P(10\% - 20\% \cup > 20\%) = 0.33 + 0.21 = 0.54$.
b. $P(B) = P(< 10\% \cup -10\% - 0\%) = 0.04 + 0.14 = 0.18$.
c. $\bar{A}$ è l'evento che il rendimento non sia maggiore del 10%.
d. $P(\bar{A}) = 0.04 + 0.14 + 0.28 = 0.46$
e. L'intersezione tra "più del 10%"" e "rendimento negativo"" è l'evento impossibile.
f. $P(A \cap B) = 0$.
g. L'unione di A e B è l'evento che il rendimento sia meno di –10%, da -10% a 0%, da 0% a 10%, da 10% a 20% e più del 20%.
h. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.54 + 0.180 = 0.72$.
i. A e B sono mutuamente esclusivi perché l'intersezione è vuota.
j. A e B non sono collettivamente esaustivi perché la probabilità della loro unione non è 1.

Esercizio 4.27

La concorrente ha in mente 3 diversi tipi di confezione E, S, L tutti ugualmente possibili. Traduzione: $P(E) = P(S) = P(L) = 1/3$.

Tre strategie di marketing tutte ugualmente possibili. Traduzione: $P(I) = P(II) = P(III) = 1/3$.

Si suppone che le scelte sul tipo di confezione e sulle strategie siano indipendenti.Traduzione: tutte le coppie di eventi $(E, I), (E, II), (E, III), (S, I), (S, II), (S, III), (L, I), (L, II), (L, III)$ sono indipendenti. La domanda riguarda la probabilità

$$P(L \cap I) =?$$

La risposta si ottiene usando la regola del prodotto con eventi indipendenti: $P(L \cap I) = P(L)P(I) = (1/3)(1/3) = 1/9$.

Esercizio 4.30

Quanti sono i modi di permutare i 6 titoli? Sono $6! = 720$. Ora supponiamo che i primi tre titoli siano A,B,C. Il numero di 6-uple che hanno i primi tre titoli uguali ad A,B,C sono formate fissando A,B,C e permutando i rimanenti 3 titoli. Il numero di queste sono $3! = 6$.

Quindi la probabilità cercata è $6/720 = 1/120 = 0.00833$.

Esercizio 4.31

Avete lo stesso problema. Comitato: `DDDDLL`

In quanto modi possibili posso costruire un sottoinsieme di 3? In $\binom{6}{3}$ modi = 20 modi. (Provare per credere).

Quanti sono i sottoinsiemi di 3 membri senza laureati? Sono $\binom{4}{3} = 4$.

Quindi la probabilità è $4/20 = 0.2$.

Esercizio 4.32

I possibili sottoinsiemi di 3 partecipanti presi da 5 sono le combinazioni di 5 oggetti di classe 3:

$$\binom{5}{3} = (5 \cdot 4 \cdot 3)/(3 \cdot 2) = 10$$

Eccole:

$$\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 3, 4\}, \{1, 3, 5\}, \{1, 4, 5\}, \{2, 3, 4\}, \{2, 3, 5\}, \{2, 4, 5\}, \{3, 4, 5\}$$

Tuttavia ognuna di queste terne si può permutare in tutti i modi. Per esempio se il podio è $\{1, 2, 3\}$, si possono avere le coppie ordinate $(1, 2, 3), (1, 3, 2), (2, 3, 1), (3, 2, 1)$ etc.

Le permutazioni dei membri della terna sono $3! = 3 \cdot 2 = 6$. Quindi il numero totale di podi è $10 \cdot 6 = 60$.

Li elenco qui sotto per chiarezza:

$$\begin{aligned}
&\{1,2,3\},\{1,2,4\},\{1,2,5\},\{1,3,2\},\{1,3,4\},\{1,3,5\},\\
&\{1,4,2\},\{1,4,3\},\{1,4,5\},\{1,5,2\},\{1,5,3\},\{1,5,4\},\\
&\{2,1,3\},\{2,1,4\},\{2,1,5\},\{2,3,1\},\{2,3,4\},\{2,3,5\},\\
&\{2,4,1\},\{2,4,3\},\{2,4,5\},\{2,5,1\},\{2,5,3\},\{2,5,4\},\\
&\{3,1,2\},\{3,1,4\},\{3,1,5\},\{3,2,1\},\{3,2,4\},\{3,2,5\},\\
&\{3,4,1\},\{3,4,2\},\{3,4,5\},\{3,5,1\},\{3,5,2\},\{3,5,4\},\\
&\{4,1,2\},\{4,1,3\},\{4,1,5\},\{4,2,1\},\{4,2,3\},\{4,2,5\},\\
&\{4,3,1\},\{4,3,2\},\{4,3,5\},\{4,5,1\},\{4,5,2\},\{4,5,3\},\\
&\{5,1,2\},\{5,1,3\},\{5,1,4\},\{5,2,1\},\{5,2,3\},\{5,2,4\},\\
&\{5,3,1\},\{5,3,2\},\{5,3,4\},\{5,4,1\},\{5,4,2\},\{5,4,3\}
\end{aligned}$$

Quindi se si tira a caso la probabilità di pescare quella vincente è 1/60.

Esercizio 4.36

a. $\binom{5}{2}\binom{6}{4} = 150$.
b. Popolazione

`MMMMM mmmmmm`

I casi possibili sono 150. Mettiamo i due fratelli nella squadra:`Mm`. Quindi la popolazione diventa

`MMMM mmmmm`

e i casi favorevoli sono il numero di modi con cui si può formare il resto della squadra:

$$\binom{4}{1}\binom{5}{3} = 40.$$

La probabilità $= 40/150 = 0.27$.

c. La probabilità che nessuno dei fratelli sia scelto si trova notando che il numero di casi favorevoli è

$$\binom{4}{2}\binom{5}{4} = 30.$$

Quindi è $30/150 = 1/5 = 0.2$.

Esercizio 4.37

Ci sono 6 fondi USA e 4 fondi INT.

a. Per fare un investimento occorre prendere 2 fondi USA e 2 INT. In quanti modi? Sono AB dove

- $A =$ Numero di possibili sottoinsiemi di 2 elementi da $6 = \binom{6}{2} = 15$
- $B =$ Numero di possibili sottoinsiemi di 2 elementi da $4 = \binom{4}{2} = 6$

Quindi $AB = 15 \times 6 = 90$.

b. Traduzione: definiamo gli eventi $C =$ Nessun fondo USA dei 2 scelti va male.
Allora siccome c'è 1 fondo USA che va male e 5 che vanno bene

$$P(C) = \frac{\binom{5}{2}}{15} = 10/15.$$

Sia poi D = Nessun fondo INT dei 2 scelti va male. Siccome c'è 1 fondo INT che va male e 3 che vanno bene:

$$P(D) = \frac{\binom{3}{2}}{6} = 3/6.$$

Sia E = Nessun fondo USA dei 2 scelti e nessun fondo INT dei 2 scelti indipendentemente va male $= C \cap D$.

Siccome sono indipendenti

$$P(E) = (10/15)(3/6) = 1/3.$$

Quindi ora si consideri che l'evento di cui si chiede la probabilità cioè F = almeno uno dei fondi scelti vada male = complementare di E avremo

$$P(F) = 1 - P(E) = 1 - 1/3 = 2/3.$$

Esercizio 4.38

A = preoccupato per il lavoro; B = preoccupato per i voti. $P(A) = 0.3$, $P(B) = 0.25$. Inoltre $P(A \cap B) = 0.2$.

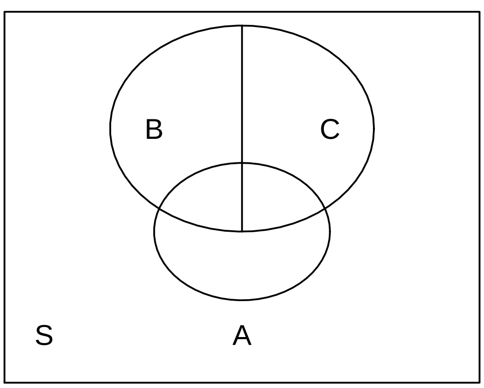
La probabilità chiesta è

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.25 - 0.2 = 0.35.$$

Esercizio 4.42

Si deve calcolare la probabilità che si verifichi almeno uno degli inconvenienti A, B, C. Questo vuol dire $P(A \cup B \cup B)$. Sappiamo che A e B sono indipendenti e quindi $P(A \cap B) = P(A)P(B)$ e anche che A e C sono indipendenti. Perciò $P(A \cap C) = P(A)P[C]$. Inoltre B e C si escludono a vicenda e quindi $P(A \cup B) = P(A) + P(B)$.

Il disegno seguente dovrebbe essere utile.



La probabilità cercata perciò è

$$\begin{aligned} P(A \cup B \cup C) &= P(B) + P[C] + P(A) - P(A \cap B) - P(A \cap C) \\ &= 0.02 + 0.01 + 0.04 - 0.02 \times 0.01 - 0.02 \times 0.04 = 0.069. \end{aligned}$$

Guardate il disegno! Quando sommo $P(A)$ conto due volte le probabilità delle intersezioni. Per questo occorre sottrarle.

ESERCIZIO DIFFICILE.

Esercizio 4.50

Sia A l'evento che il problema capita il lunedì. Sia B l'evento che il problema succeda durante l'ultima ora. Quindi $P(A) = 0.3$ e $P(B) = 0.2$. Inoltre siccome il 4% dei problemi avviene di lunedì e all'ultima ora vuol dire che $P(A \cap B) = 0.04$.

a. Chiede la probabilità che sapendo che un problema si verifica di lunedì questo NON si verifichi alla fine del turno, cioè
$$P(\bar{B}|A) = \frac{P(\bar{B} \cap A)}{P(A)}$$
Il numeratore è la probabilità della differenza tra gli eventi A e B. Questa si trova (fare un disegno) con
$$P(\bar{B} \cap A) = P(A) - P(A \cap B) = 0.3 - 0.04 = 0.26.$$
Quindi $P(\bar{B}|A) = 0.26/0.3 = 0.867$.

b. A e B sono indipendenti se e solo se $P(A \cap B) = P(A)P(B)$ e questo è chiaramente falso.

Esercizio 4.52

A = Chi risponde alla telefonata diventa un nuovo cliente B = Il rispondente si era avvalso l'anno prima dei servizi della concorrenza

- $P(A) = 0.15$
- $P(B|A) = 0.8$
- $P(B) = 0.6$

$P(A|B) = ?$

Per la formula di Bayes

$$P(A|B) = P(B|A)P(A)/P(B) = 0.8 \cdot 0.15/0.6 = 0.2$$

Esercizio 4.74

Definite l'evento $G =$ gli utili crescono.

Inoltre consiederate i tre eventi H, L e S definiti come segue.

$H =$ i tassi del prossimo anno superano quelli correnti di più dell'1%

$L =$ i tassi del prossimo anno sono inferiori a quelli correnti di più dell'1%

$S =$ i tassi del prossimo anno si scostano da quelli correnti di meno dell'1%.

Per costruzione H, L, S sono una partizione dello spazio campionario (sono eventi necessari e incompatibili).

Allora si sa che $P(G|H) = 0.1$, $P(G|L) = 0.8$ e $P(G|S) = 0.5$. Inoltre si sa anche che

$$P(H) = 0.25, \quad P(L) = 0.15, \quad P(S) = 0.6.$$

Notate che la somma fa 1.

a. Si chiede

$$P(G \cap H) = P(G|H)P(H) = (0.1)(0.25) = 0.025$$

b. Si chiede

$$P(G) = P(G|H)P(H) + P(G|L)P(L) + P(G|S)P(S)$$

Questa è la formula delle probabilità totali. Si ottiene

$$P(G) = 0.025 + (0.8)(0.15) + (0.5)(0.6) = 0.445$$

c. Chiede

$$P(L|G) = P(G|L)P(L)/P(G) = (0.8)(0.15)/0.445 = 0.2697.$$

Questa è la formula di Bayes.

Esercizio 4.76

L'esercizio è formulato in modo un po' confuso. Se si legge attentamente alla fine si capisce che:

- L'esperimento consiste nell'estrazione casuale di uno studente universitario.
- Lo studente può essere bravo (B) , medio (M), scarso (S).

Siccome lo studente è bravo se si colloca tra il 25% dei migliori P(B) = 0.25 Siccome lo studente è scarso se si colloca tra il 25% dei peggiori P(S) = 0.25 Evidentemente P(M) = 0.5.

- Lo studente può essere stato ottimo (O) alle superiori.

Siccome lo studente è ottimo se si collocava tra il 10% dei migliori P(O) = 0.1.

- Ci viene detto che

P(O | B) = 0.7 P(O | M) = 0.5 P(O | S) = 0.2

Quindi possiamo impostare la tabella seguente

```
          O    non O  Tot

B     0.7     0.3     1
M     0.5     0.5     1
S     0.2     0.8     1

Tot   0.1     0.9     1
```

A questo punto vengono le domande.

a. Chiede P(O). Usando la formula delle probabilità totali si ha

$$P(O) = P(O|B)P(B) + P(O|C)P(C) + P(O|S)P(S) = (0.7)(0.25) + (0.5)(0.5) + (0.2)(0.25) = 0.475.$$

Esercizio 4.78

(Sono esercizi normali che potete trovare nel compito).

$$P(M_1) = 0.4, \quad P(M_2) = 0.6$$

Sono le probabilità che le caramelle siano prodotte dalla macchina 1 e 2.

- $P(D \mid M_1) = 0.1$
- $P(D \mid M_2) = 0.$

È la probabilità di produrre scatole difettose (sapore alterato) sapendo che il pezzo proviene dalla macchina 1 o 2.

La macchina 2 non produce pezzi difettosi e quindi la probabilità è 0.

$$P(M_1 \mid nonD) = ?$$

è la probabilità che la scatola provenga dalla macchina 1 sapendo che il sapore non è alterato.

È una applicazione della regola di Bayes.

$$P(M_1 \mid nonD) = \frac{P(nonD \mid M_1)P(M_1)}{P(nonD)}$$

$$P(nonD \mid M_1)P(M_1) = 0.36$$
$$P(nonD) = P(nonD \mid M_1)P(M_1) + P(nonD \mid M_2)P(M_2) = 0.9 \cdot 0.4 + 1 \cdot 0.6 = 0.96$$

Quindi la soluzione è $0.36/0.96 = 0.375$.

Esercizio 4.82

Notare che A_1 e A_2 sono complementari cioè $P(A_2) = 1 - P(A_1)$.

Stessa cosa per B_1 e B_2. Quindi per la formula di Bayes:

$$P[A_1 \mid B_2] = P[B_2|A_1]P[A_1]/P[B_2] = (1 - P[B_1|A_1])P[A_1]/P[B_2]$$

Perché B_1 è complementare di B_2 !

Ora $P[B_1] = P[B_1|A_1]P[A_1] + P[B_1|A_2]P[A_2]$ per il teorema della probabilità totale.

Quindi

$$P[B_1] = 0.4 \cdot 0.5 + 0.7 \cdot 0.5 = 0.55$$

e quindi $P[B_2] = 0.45$.

Allora $P[A_1|B_2] = (1 - 0.4) \cdot 0.5/0.45 = 0.66667$.

Esercizio DIFFICILE.

Esercizio 4.85

$A =$ Il prof. riceve il materiale $B =$ il prof. adotta il libro

Sappiamo che

$$P(A) = 0.8, \quad P(B|A) = 0.3 \quad P(B|\bar{A}) = 0.1$$

Perciò si chiede $P(A|B$ che si trova con la formula di Bayes

$$\begin{aligned} P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid \bar{A})P(\bar{A})} \\ &= \frac{(0.3)(0.8)}{(0.3)(0.8) + (0.1)(1 - 0.8)} = 0.923 \end{aligned}$$

Esercizio 4.86

- A_1 = molto meglio; A_2 = molto peggio; A_3 = nella media.
- B = ottimo acquisto

$$P[A_1] = 0.25; P[A_2] = 0.25; P[A_3] = 0.5$$
$$P[B|A_1] = 0.4; P[B|A_2] = 0.1; P[B|A_3] = 0.2$$
$$P[A_1|B] = (.25 \cdot 0.4)/(.25 \cdot 0.4 + .25 \cdot .1 + .5 \cdot .2) = 0.444444.$$

Esercizio 4.95

W = incidente causato da condizioni atmosferiche. BI = l'incidente ha provocato feriti. Abbiamo che

$$P(W) = 0.3 \quad P(BI) = 0.2 \quad P(W|BI) = 0.4$$

a. $P(W \cap BI) = P(W|BI)P(BI) = (0.4)(0.2) = 0.08.$

b. No, infatti $P(W \cap BI) = 0.08 \neq 0.06 = P(W)P(B).$

c. $P(BI \mid W) = P(W \cap BI)/P(W) = 0.08/0.3 = 0.267.$

d. Abbiamo

$$\begin{aligned} P(\bar{W} \cap \overline{BI}) &= P(\overline{W \cup BI}) = 1 - P(W \cup BI) \\ &= 1P(W)P(BI) + P(W \cap BI) = 1 - 0.3 - 0.2 + 0.08 = 0.58 \end{aligned}$$

Esercizio 4.96

Sia G = arrivo cavi grossi e F = arrivo cavi fini.

Si sa che $P(G \cup F) = 0.8$ e inoltre $P(G \mid F) = 0.4$ mentre $P(F \mid G) = 0.6$.

Perciò se si pone

$$P(G) = x; P(F) = y; P(G \cap F) = z$$

risulta

$$\begin{aligned} 0.8 &= P(G \cap F) = x + y - z \\ 0.4 &= P(G|F) = z/y \\ 0.6 &= P(F|G) = z/x \end{aligned}$$

È un sistema di tre equazioni in 3 incognite che va risolto e quindi si ottiene

$$x = 0.4211, \quad y = 0.6316, \quad z = 0.2526.$$

Queste sono le soluzioni dei quesiti a. b. c.

E' TROPPO DIFFICILE PER VOI!

Esercizio 4.101

Definiamo $N =$ turno di notte, $F = femmina$, $M = maschio$, $FP =$ favorevole al piano.
Sappiamo che

$$\begin{aligned} P(N) = 0.5, \quad P(F) = 0.3, \quad P(M) = 0.7, \quad P(F|N) = 0.2, \\ P(M|N) = 0.8, \quad P(FP|N) = 0.65, P(FP|F) = 0.4. \end{aligned}$$

Perciò

$$\begin{aligned} P(N \cap F) &= (0.5)(0.2) = 0.1, \\ P(FP) &= P(FP|M)P(M) + P(FP|F)P(F) = (0.5)(0.7) + (0.4)(0.3) = 0.47 \\ P(N \cap FP) &= P(FP|N)P(N) = (0.65)(0.5) = 0.325 \end{aligned}$$

a.
$$P(FP \cap F) = P(FP|F)P(F) = (0.4)(0.3) = 0.12$$

b.
$$P(N \cup F) = P(N) + P(F)P(N \cap F) = 0.5 + 0.3 - 0.1 = 0.7$$

c. No, verifica che
$$P(N \cap F) = 0.1 \neq 0.15 = P(N)P(F).$$

d.
$$P(N|F) = P(N \cap F)/P(F) = 0.1/0.3 = 0.3333.$$

e. Abbiamo

$$\begin{aligned} P(\bar{N} \cap \overline{FP}) &= 1 - P(N \cup FP) = 1 - [P(N) + P(FP) - P(N \cap FP)] \\ &= 1 - 0.5 - 0.47 + 0.325 = 0.355. \end{aligned}$$

Esercizio 4.102

a. C'è un gruppo di 16 persone. Devo fare una giuria di 12. Quante scelte posso fare? È il numero di sottoinsiemi di 12 elementi da un insieme di 16. Quindi calcolo le combinazioni di 16 elementi di classe $12 = \binom{16}{12}$. Quindi calcolo

$$\binom{16}{12} = (16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5)/(12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)$$

Nota che sono entrambi prodotti di 12 numeri: semplifico

$$(16 \cdot 15 \cdot 14 \cdot 13)/(4 \cdot 3 \cdot 2 \cdot 1) = 1820$$

b. Se pesco a caso 12 persone qual è la probabilità che ci siano più uomini che donne? Facciamo un esempio più semplice: qual è la probabilità che ci siano 7 uomini? La chiamo $P(7)$.

Il numero di casi possibili è 1820. Il numero di casi favorevoli è il numero di sottoinsiemi di 12 elementi con 7 uomini e 5 donne presi da un insieme di 16. In quanti modi posso prendere 7 uomini da un gruppo di 8? È

$$\binom{8}{7} = (8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2)/(7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) = 8.$$

In quanti modi posso prendere 5 donne da un gruppo di 8? È

$$\binom{8}{5} = (8 \cdot 7 \cdot 6 \cdot 5 \cdot 4)/(5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) = 56.$$

Quindi per ognuno degli 8 modi di prendere 7 maschi ci sono 56 modi di accostargli 5 femmine. Quindi il numero di casi favorevoli è $8 \cdot 56 = 448$.

La probabilità cercata è

$$P(7) = \binom{8}{7}\binom{8}{5}/1820 = 448/1820 = 0.246.$$

Ora la probabilità che ci siano più uomini che donne include anche la probabilità che la giuria si composta da 8 uomini: $P(8)$. Usando lo stesso ragionamento calcolo

$$P(8) = \binom{8}{8}\binom{8}{4}/1820 = 1 \cdot 70/1820 = 0.0385.$$

Quindi la risposta è $P(7) + P(8) = 0.246 + 0.0385 = 0.284$.

Esercizio 4.104

Sappiamo che se A = essere selezionato (e sottoposto alla cura), B = guarigione

$$\begin{aligned} P(A) &= 10/100 = 0.10 \\ P(B|A) &= 0.75 \\ P(B|nonA) &= 0.50 \end{aligned}$$

Quindi

a) $P(B \cap A) = P(A)P(B|A) = 0.10 \cdot 0.75 = 0.075$.

b) $P(A|B) = P(A \cap B)/P(B) = P(B \cap A)/P(B) = 0.075/P(B)$. Ma

$$P(B) = P(B|A)P(A) + P(B|nonA)P(nonA) = 0.075 + 0.50 \cdot (1 - P(A)) = 0.075 + 0.5 \cdot 0.9 = 0.525.$$

Quindi $P(A|B) = 0.075/P(B) = 0.075/0.525 = 0.1429$.

c) Diciamo che chiamiamo 10 pazienti specifici in un certo modo: Tizio, Caio, Sempronio ... Giulia.

Qual è la probabilità di pescare dai 100 pazienti proprio i precedenti 10?

Il numero di casi possibili è $\binom{100}{10}$. Il numero di casi favorevoli sono 1!

Quindi abbiamo che la probabilità è

$$1/\binom{100}{10} = 10!90!/100! = 1/17310309456440.$$

Esercizio 5.20

d. Quando dice "si scelgono a caso 2 scatole" vuol dire che il numero di graffette X_1 della prima scatola e X_2 della seconda hanno la distribuzione di probabilità uguale a quella data e che sono indipendenti fra loro. La probabilità la prima scatola contenga come minimo 50 graffette è

$$P(X_1 \geq 50) = 0.29 + 0.2 + 0.1 + 0.03 = 0.62.$$

Chiamo questo evento $A_1 = X_1 \geq 50$. Ovviamente anche per la seconda scatola $P(X_2 \geq 50) = 0.62$. Chiamo questo secondo evento $A_2 = X_2 \geq 50$. Ora chiede infine la probabilità che almeno una delle scatole contenga 50 o più graffette cioè

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Ora siccome le estrazioni delle due scatole sono indipendenti A_1 e A_2 sono due eventi indipendenti. Perciò abbiamo

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1)P(A_2)$$

e dunque

$$P(A_1 \cup A_2) = 0.62 + 0.62 - 0.62 \times 0.62 = 0.8556.$$

e. Diciamo che $Y = 16 + 2X$ centesimi sia il costo di produzione e perciò l'utile è $U = 150 - Y = 150 - 16 - 2X = 134 - 2X$ centesimi. Il valor medio e la deviazione standard dell'utile sono perciò

$$\begin{aligned} E(U) &= 134 - 2E(X) \\ \mathrm{var}(U) &= 4\mathrm{var}(X) \end{aligned}$$

Ora non resta che calcolare la media e la varianza di X. Un pò di conti permettono di trovare

$$E(X) = 49.9 \qquad \mathrm{var}(X) = 1.3964^2$$

Quindi

$$E(X) = 134 - 2 \times 49.9 = 34.2 \text{ centesimi } (\$0.34).$$

e

$$\sigma(U) = 2 \times 1.3964 = 2.7928 \text{ centesimi } (\$0.0279).$$

Esercizio 5.22

Un lotto contiene il 10% di pezzi difettosi. Si scelgono a caso due pezzi e si controllano. X = numero di pezzi difettosi su due a. Calcolare la distribuzione di probabilità.

x	$P(x)$
0	0.81
1	0.18
2	0.01

Infatti $X \sim Bin(n = 2, p = 0.1)$

$$\begin{aligned} P(X = 0) &= 0.9 \cdot 0.9 = 0.81 \\ P(X = 1) &= 2 \cdot 0.9 \cdot 0.1 = 0.18 \\ P(X = 2) &= 0.1 \cdot 0.1 = 0.01 \end{aligned}$$

b. Confezione di 20 pezzi con due pezzi difettosi. Si scelgono a caso due pezzi e si controllano.

Y = numero di pezzi difettosi su due

y	$P(y)$
0	0.8053
1	0.1895
2	0.0053

Infatti

$$P(Y=0) = \frac{18}{20}\frac{17}{19} = 0.8053$$
$$P(Y=1) = \frac{2}{20}\frac{18}{20} + \frac{18}{20}\frac{2}{19} = 0.1895$$
$$P(Y=2) = \frac{2}{20}\frac{1}{19} = 0.0053$$

c. Abbiamo

$$E(X) = 0 \cdot 0.81 + 1 \cdot 0.18 + 2 \cdot 0.01 = 0.2$$
$$\text{var}(X) = 0^2 \cdot 0.81 + 1^2 \cdot 0.18 + 2^2 \cdot 0.01 - 0.2^2 = 0.18$$

d. Abbiamo

$$E(Y) = \sum yp(y) = 0 \cdot 0.8053 + 1 \cdot 0.1895 + 2 \cdot 0.0053 = 0.2001$$
$$\text{var}(Y) = \sum y^2 p(y) - \mu_y^2 = 0^2 \cdot 0.8053 + 1^2 \cdot 0.1895 + 2^2 \cdot 0.0053 - 0.2001^2 = 0.1707$$

Esercizio 5.27

È LUNGO e DIFFICILE!!

P = profitto X = numero di richieste S = numero di giornali in magazzino

$$P = \begin{cases} 0.2S - 0.05(XS) & \text{se } X > S \\ 0.2X & \text{se } X = S \\ 0.2X - 0.7(SX) & \text{se } X < S \end{cases}$$

Profitto per ogni combinazione di S e X.

S, X	0	1	2	3	4	5
0	0.00	−0.05	−0.10	−0.15	−0.20	−0.25
1	−0.70	0.20	0.15	0.10	0.05	0.00
2	−1.40	−0.50	0.40	0.35	0.30	0.25
3	−2.10	−1.20	−0.30	0.60	0.55	0.50
4	−2.80	−1.90	−1.00	−0.10	0.80	0.75
5	−3.50	−2.60	−1.70	−0.80	0.10	1.00

Quindi calcolano i valori medi per ogni distribuzione condizionata.

$$
\begin{aligned}
E[P|S=0] &= 0 + (-.05)(.16) + (-.1)(.18) + (-.15)(.32) + \\
&+ (-.2)(.14) + (-.25)(.08) = -.122 \\
E[P|S=1] &= (-.7)(.12) + (.2)(.16) + (.15)(.18) + \\
&+ (.1)(.32) + (.05)(.14) = .014 \\
E[P|S=2] &= (-1.4)(.12) + (-.5)(.16) + (.4)(.18) + \\
&+ (.35)(.32) + (.3)(.14) + (.25)(.08) = -.002 \\
E[P|S=]3 &= (-2.1)(.12) + (-1.2)(.16) + (-.3)(.18) + \\
&+ (.6)(.32) + (.55)(.14) + (.5)(.08) = -.189 \\
E[P|S=4] &= (-2.8)(.12) + (-1.9)(.16) + (-1)(.18) + \\
&+ (-.1)(.32) + (.8)(.14) + (.75)(.08) = -.68 \\
E[P|S=5] &= (-3.5)(.12) + (-2.6)(.16) + (-1.7)(.18) + \\
&+ (-.8)(.32) + (.1)(.14) + (1)(.08) = -1.304
\end{aligned}
$$

Conclusione: il proprietario massimizza il profitto atteso ordinando 1 giornale.

Esercizio 5.40

La distribuzione del numero di partite vinte è binomiale con $n = 5$ e $p = 0.4$ Munendosi di pazienza oppure di una buona calcolatrice si ottiene la funzione di massa di probabilità e la funzione di ripartizione.

x	p(x)
0	0.07776
1	0.25920
2	0.34560
3	0.23040
4	0.07680
5	0.01024
Tot	1.00000

x	F(x)
0	0.07776
1	0.33696
2	0.68256
3	0.91296
4	0.98976
5	1.00000

a. $P(X = 5) = 0.01024$
b. $P(X \geq 3) = 0.23040 + 0.07680 + 0.01024 = 0.31744$.
c. Se vincono la prima partita restano 4 partite per cui il numero di partite vinte Y è Binomiale con $n = 4$ e $p = 0.4$.

La distribuzione è

y	p(y)
0	0.1296
1	0.3456
2	0.3456
3	0.1536
4	0.0256

Ora la probabilità di vincere la maggioranza delle 5 partite è la probabilità di vincerne almeno 2 delle 4 restanti cioè

$$1 - P(Y = 0) - P(Y = 1) = 0.5248$$

d. E' la media delle partite vinte cioè

$$E(X) = np = 5 \times 0.4 = 2.$$

e. Dato che i Clubs hanno vinto la prima il numero medio di partite vinte è $1 + E(Y) = 1 + 4 \times 0.4 = 2.6$.

E' un esercizio difficilino.

Esercizio 5.50

La regola è basata sulla formula

$$\frac{\binom{S}{x}\binom{N-S}{n-x}}{\binom{N}{n}}$$

dove N = numero di elementi della popolazione finita, S = numero di successi nella popolazione, n = dimensione del campione, x = numero di successi nel campione. Quindi

$$\frac{\binom{25}{5}\binom{25}{7}}{\binom{50}{12}} = 0.2104$$

Esercizio 5.54

Nella popolazione ci sono $N = 400$ unità di cui $S = 200$ successi. Estraggo un campione casuale senza ripetizione di dimensione $n = 15$. Il numero di possibili campioni è

$$\binom{400}{15} = \frac{400 \cdot 399 \cdot 398 \cdot \cdots \cdot 386}{15!} = 629498843449501998439382160$$

Evidentemente non potete calcolarlo a mano!

Il numero di possibili campioni con 8 successi e 6 insuccessi sono

$$\binom{200}{8}\binom{200}{7} = 125840388797424048232485000$$

Quindi la probabilità chiesta è il rapporto di questi due valori che risulta

$$\frac{125840388797424048232485000}{629498843449501998439382160} = 0.1999.$$

ESERCIZIO IMPOSSIBILE DA FARE A MANO!

Esercizio 5.58

Se X è il numero di richieste inoltrate da immigrati, X è una variabile ipergeometrica con $N = 10$, numero di successi (il numero di immigrati nella popolazione) $S = 5$ e numerosità del campione $n = 6$.

Qui le richieste approvate (per assurdo) sono state estratte a caso. Con molta pazienza e una calcolatrice si ottiene la funzione di massa di probabilità e la funzione di ripartizione.

```
x    p(x)       F(x)
0 0.00000000 0.00000000
1 0.02380952 0.02380952
2 0.23809524 0.26190476
3 0.47619048 0.73809524
4 0.23809524 0.97619048
5 0.02380952 1.00000000
```

La probabilità che meno della metà delle richieste APPROVATE sia stata inoltrata da immigrati è la probabilità che su 6 richieste approvate al massimo 2 siano state inoltrate da immigrati cioè

$$P(X \leq 2) = 0.2619.$$

ESERCIZIO DIFFICILE!

Esercizio 5.74

Le medie sono $\mu_X = 1.5$ e $\mu_Y = 0.55$. La covarianza è

$$\begin{aligned}\sigma_{XY} &= \sum_x \sum_y xyP(x,y) - \mu_x\mu_y \\ &= 0.3 \times 1 + 0.25 \times 2 - 1.5 \times 0.55 = -0.025.\end{aligned}$$

La varianza di X è

$$\sigma_X^2 = \sum x^2 P(x) - \mu_x^2 = 1 \times 0.5 + 4 \times 0.5 - 1.5^2 = 0.25$$

Analogamente la varianza di Y è $\sigma^2 = 0.2475$. Perciò il coefficiente di correlazione è

$$\rho_{XY} = \frac{-0.025}{\sqrt{0.25 \times 0.2475}} = -0.1005.$$

Esercizio 5.80

a. Distribuzione marginale di X: $P(X = 1) = 0.4, P(X = 2) = 0.6$. Distribuzione marginale di Y:$P(Y = 0) = 0.6$, $PY = 1) = 0.4$.

b. Media e Deviazione standard di X:

$$\begin{aligned}E(X) &= 1(0.4) + 2(0.6) = 0.4 + 1.2 = 1.6 \\ \text{var}(X) &= 1^2(0.4) + 2^2(0.6) - 1.6^2 = 0.24 \\ E(Y) &= 0(0.6) + 1(0.4) = 0.4 \\ \text{var}(X) &= 0^2(0.6) + 1^2(0.4) - 0.4^2 = 0.24\end{aligned}$$

Covarianza tra X e Y

$$\sigma_{XY} = 0(1)(0) + 0(2)(0.6) + 1(1)(0.4) + 1(2)(0) - 1.6(0.4) = -0.24$$

Coefficiente di correlazione

$$\rho_{XY} = \frac{(-0.24)}{\sqrt{0.24(0.24)}} = -1.$$

c. Media di W

$$E(W) = 2\mu_X - 4\mu_Y = 2(1.6) - 4(0.4) = 1.6$$

Varianza di W

$$\begin{aligned}\sigma_W^2 &= 4\sigma_X^2 + 16\sigma_Y^2 - 2(2)(4)\sigma_{XY}^2 \\ &= 4(0.24) + 16(0.24) - 2(2)(4)(-0.24) = 8.64.\end{aligned}$$

Esercizio 5.99

a. La probabilità di vincere almeno 4 partite su 7 è (supponendo l'indipendenza delle partite e la probabilità di vittoria costante) Binomiale

$$P(X \geq 4) = P(X = 4 \text{ o } 5 \text{ o } 6 \text{ o } 7)$$

quindi è (forza con la calcolatrice!)

$$\binom{7}{4} 0.6^4 0.4^3 + \binom{7}{5} 0.6^5 0.4^2 + \binom{7}{6} 0.6^6 0.4^1 + 0.6^7 = 0.71021$$

b. La probabilità che la settima gara sia decisiva è la probabilità che si siano vinte esattamente 3 partite sulle prime 6 e cioè

$$P(3 \text{ successi su } 6 \text{ partite}) = \binom{6}{3} 0.6^3 0.4^3 = 0.27648$$

c.i. Una volta che le prime 4 partite sono finite in parità la probabilità che A vinca è che vinca almeno 2 partite sulle restanti 3:

$$\binom{3}{2} 0.6^2 0.4^1 + \binom{3}{3} 0.6^3 0.4^0 = 0.648.$$

c.ii La probabilità che sia necessaria una 7ma partita per determinare il vincitore è la probabilità che sulle prossime 2 partite A ne vinca 1:

$$\binom{2}{1} 0.6^1 0.4^1 = 0.48.$$

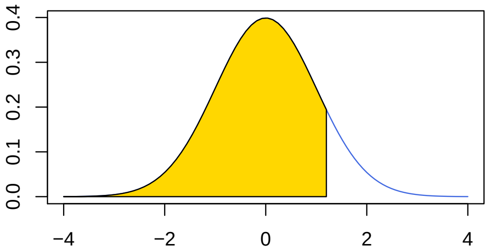
DIFFICILE.

Capitolo 6. Variabili casuali continue (la normale)

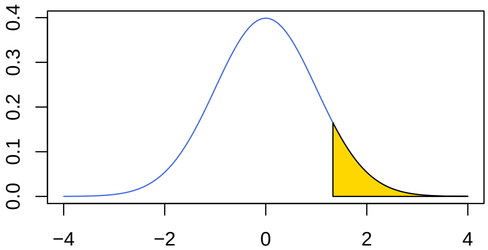
Esercizio 6.17

- È un esercizio UTILE e FONDAMENTALE!! DISEGNATE LE FIGURE!
- Si usano le tavole della normale standard e si sfrutta la simmetria.

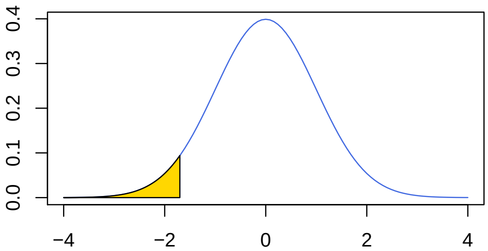
a. $P(Z < 1.20) = F(1.20) = 0.8849$



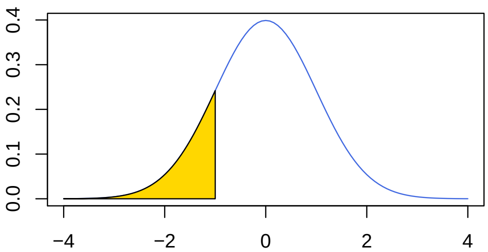
b. $P(Z > 1.33) = 1 - P(Z \leq 1.33) = 1 - F(1.33) = 1 - 0.9082 = 0.0918$



c. $P(Z < -1.70) = P(Z \geq 1.70) = 1 - P(Z < 1.70) = 1 - F(1.70) = 1 - 0.9554 = 0.0446$

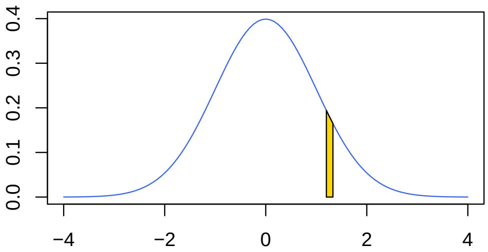


d. $P(Z > -1) = P(Z < 1) = F(1) = 0.8413$

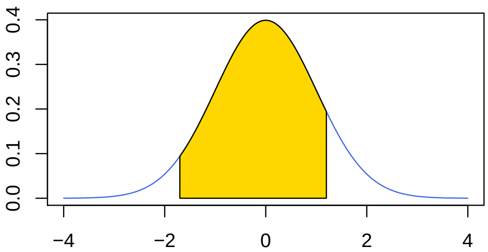


e. $P(a < Z < b) = P(Z < b) - P(Z < a)$

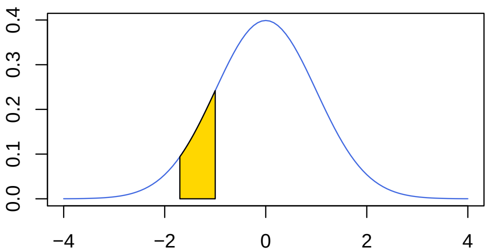
$$\begin{aligned} P(1.20 < Z < 1.33) &= P(Z < 1.33) - P(Z < 1.20) \\ &= F(1.33) - F(1.20) \\ &= 0.9082 - 0.8849 = 0.0233 \end{aligned}$$



f. $P(-1.70 < Z < 1.20) = F(1.20) - [1 - F(1.70)] = 0.8849 - (1 - 0.9554) = 0.8403$



g. $P(-1.70 < Z < -1) = P(1 < Z < 1.70) = F(1.70) - F(1) = 0.9554 - 0.8413 = 0.1141.$



NOTA BENE: in questi calcoli non importa se viene usato $<$ o $\leq$! Infatti per esempio

$$P(Z \leq 1) = P(Z < 1) + P(Z = 1) = P(Z < 1) + 0$$

perché la normale è una variabile continua.

Esercizio 6.18

a. Vai sulle tavole e guardi il valore più vicino a 0.70. Trovi

$$\begin{aligned} F(z) &= 0.6985, \quad z = 0.52 \\ F(z) &= 0.7019, \quad z = 0.53 \end{aligned}$$

Puoi fare in due modi: 1) Fai la semisomma $(0.52 + 0.53)/2 = 0.525$ oppure 2) guardi quale tra i due valori di probabilità è più vicino a 0.70.

In questo caso è $z = 0.52$. Infatti 0.6985 è più vicino a 0.70 di 0.7019.

b. Per trovare z tale che $P(Z < z) = 0.25$ non si può guardare sulla Tavola 1 perché questa per economia di spazio fornisce solo probabilità tra 0.5 e 0.999.

Quindi si ragiona in questo modo. Si risolve il problema simmetrico

$$P(Z < z) = 0.75$$

dove $0.75 = 1 - 0.25$. Si trova col metodo del punto precedente che $z = 0.67$. Come si vede 0.67 è il terzo quartile della normale standard.

Il valore cercato che lascia a sinistra un"area 0.25 deve invece essere $z = -0.67$ perché è simmetrico rispetto alla media 0. Infatti

$$P(Z < -0.67) = 1 - P(Z < 0.67) = 1 - 0.75 = 0.25.$$

Si nota anche che $z = -0.67$ è il primo quartile della normale standard.

c. Per trovare $P(Z > z) = 0.20$ si fa così: si osserva che se z soddisfa la relazione precedente allora

$$P(Z < z) = 0.80$$

Quindi si procede come per il caso a. e si trova $z = 0.84$.

d. Infine se dobbiamo trovare $P(Z > z) = 0.60$ allora sarà

$$P(Z < z) = 0.4$$

Siccome la probabilità è più piccola di 0.5 si procede come nel caso b. Cioè prima si calcola

$$P(Z < z) = 0.6$$

e si trova $z = 0.25$. Poi si trova il valore di z giusto cambiando di segno. Cioè

$$P(Z < -0.25) = 0.4$$

e anche

$$P(Z > -0.25) = 0.6$$

Ognuna di queste soluzioni si può verificare alla fine. Per esempio nel caso d. si verifica che

$$P(Z < -0.25) = 1 - P(Z < 0.25) = 1 - 0.5987 \approx 0.4.$$

Naturalmente le soluzioni trovate sono sempre approssimate perché usiamo la Tavola 1 in modo inverso.

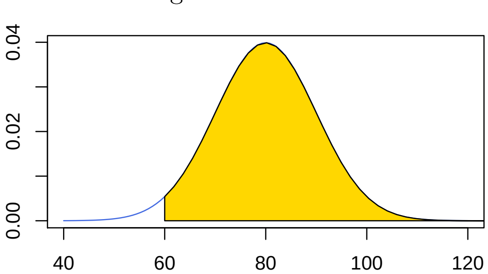
Esercizio 6.20

$X \sim N(80, \sigma = 10)$

a. Abbiamo

$$P(X > 60) = 1 - P(X < 60) = 1 - P(Z < (60 - 80)/10) = 1 - P(Z < -2) = P(Z < 2) = 0.9772.$$

Fate un grafico della normale standard!!



b. Abbiamo

$$P(72 < X < 82) = P(X < 82) - P(X < 72)$$

Quindi si ottiene

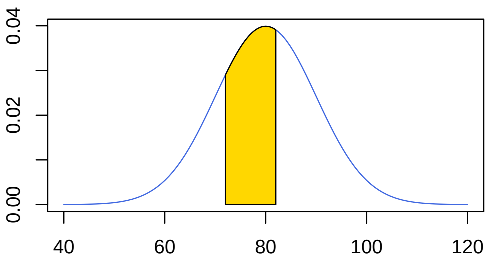
$$P(Z < (82 - 80)/10) - P(Z < (72 - 80)/10) = P(Z < 0.2) - P(Z < -0.8)$$

Dalle tavole

$$P(Z < 0.2) = 0.5793$$

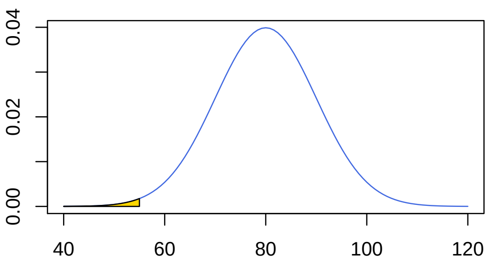
$$P(Z < -0.8) = 1 - P(Z < 0.8) = 1 - 0.7881$$

e quindi il risultato è $0.5793 - 1 - 0.7881 = 0.3674$.



c. Abbiamo

$$P(X < 55) = P(Z < (55 - 80)/10) = P(Z < -2.50) = 1 - P(Z < 2.50) = 0.0062.$$



d. Come trovare x tale che $P(X > x) = 0.1$?

Ora, si sa che $P(Z > 1.282) = 0.1$ per la normale standard Z (guardate le tavole della t di Student ultima riga).

Quindi, poiché $X = 80 + 10Z$, si può scrivere

$$P(80 + 10Z > 80 + 10 \cdot 1.282) = 0.1$$

cioè

$$P(X < 92.82) = 0.1$$

e quindi la risposta è $x = 92.82$.

e. Notare che se la probabilità che X assuma valori all'esterno è 0.08 allora in ciascuna coda deve essere 0.04. Perciò prima si cerca per la normale $N(0, 1)$

$$P(Z > z) = 0.04$$

Questo non si trova sulle tavole della t di Student. Occorre guardare la tavola 1 della normale all'inverso, in corrispondenza di $P(Z < z) = 0.96$ e si vede che $z = 1.75$. Questo valore è chiamato $z_{\alpha/2}$ dove $\alpha = 0.08$.

Perciò l'intervallo $(-1.75, 1.75)$ è centrato nella media e ha le caratteristiche richieste per la normale standard.

Ma qui abbiamo la normale $N(80, \sigma = 10)$ e quindi l'intervallo richiesto si ottiene applicando l'inverso della trasformazione di standardizzazione, cioè:

$$\text{Estremo inferiore } = 80 - 1.75 \cdot 10 = 62.5, \quad \text{Estremo superiore } = 80 + 1.75 \cdot 10 = 97.5$$

Notare la somiglianza con gli intervalli di confidenza (Capitolo 9).

Esercizio 6.23

c. Qual è il numero di unità vendute che ha probabilità 0.1 di essere superato? Se $X \sim N(1200, 100^2)$ sono le vendite, si deve calcolare il valore d tale che

$$P(X > d) = 0.1$$

Devi usare le tavole all'inverso. Quindi standardizzo ambo i membri:

$$P(Z > \frac{d - 1200}{100}) = 0.1$$

Sulle tavole trovo che

$$P(Z \leq 1.28) \simeq 0.9$$

e quindi $P(Z > 1.28) \simeq 0.1$. Perciò

$$\frac{d - 1200}{100} = 1.28$$

e allora il valore cercato d è

$$d = 1200 + 1.28 \times 100 = 1328.$$

Esercizio 6.28

Se $X \sim N(\mu, \sigma)$ è il punteggio al test si chiede

$$p = P(X > \mu + 1.5\sigma)$$

Quindi standardizzando ambo i membri rispetto alla media e alla deviazione standard (incognite) di X si ottiene

$$p = P(Z > [(\mu + 1.5\sigma) - \mu]/\sigma) = P(Z > 1.5)$$

e quindi

$$p = 1 - P(Z \leq 1.5) = 0.0668.$$

Esercizio 6.30

Sia X lo *share* televisivo e supponiamo che abbia una distribuzione di probabilità normale di media μ e deviazione standard σ. Si giudica che

$$P(X > 17.8) = 0.25, \text{ e } P(X > 19.2) = 0.15.$$

Per la normale standard si trova consultando le tavole:

$$P(Z > 0.67) = 0.25 \text{ e } P(Z > 1.03) = 0.15.$$

(Guardate $F(z)$ e trovare lo z corrispondente.)

Perciò i valori 17.8 e 19.2 si possono scrivere come segue

$$\begin{aligned} \mu + \sigma 0.67 &= 17.8 \\ \mu + \sigma 1.03 &= 19.2 \end{aligned}$$

Questo è un sistema lineare di due equazioni in due incognite che si può risolvere e fornisce come soluzioni

$$\mu = 15.265, \sigma = 3.7838.$$

È UN ESERCIZIO DIFFICILE!

Esercizio 6.34

a. Traduzione: se scelgo a caso 1 studente qual è la probabilità che il tempo x che è superato dal 90% dei tempi nella popolazione? Cioè:

$$P(X > x) = 0.9$$

Allora, si standardizza rispetto alla media e varianza di X, e si ottiene

$$P(Z > (x - 150)/40) = 0.9$$

Intanto cerchiamo k tale che per $Z \sim N(0, 1)$,

$$P(Z > k) = 0.9$$

Sulla Tavola 2 della t di Student all'ultima riga (quella della normale standard) si trova

$$P(Z > 0.1) = 1.282$$

Quindi per simmetria avremo

$$P(Z > -1.282) = 0.9$$

Quindi il valore x cercato si trova risolvendo l'equazione

$$(x - 150)/40 = -1.282$$

ossia $x = 150 + (40)(-1.282) = 98.72$.

b. È simile: trovare

$$P(X < x) = 0.8$$

Cioè

$$P(Z < k) = P(Z < (x - 150)/40) = 0.8$$

Questa volta si va sulla Tavola 1 e si cerca il valore più vicino a 0.8. Si vede che è

$$P(Z < 0.84) \simeq 0.8$$

quindi x soddisfa

$$(x - 150)/40 = 0.84$$

che si risolve ottenendo $x = 150 + 40(0.84) = 183.6$.

c. Tempo = X. Se seleziono due studenti in modo casuale i due tempi dei due studenti sono X_1 e X_2 hanno la stessa distribuzione normale e sono indipendenti. Chiamo successo per il primo studente se dedica 2 ore o più al lavoro. Lo stesso per il secondo studente.

La probabilità di successo è $p = P(X \geq 120min) = P(Z \geq (120 - 150)/40)$ cioè $1 - P(Z < -0.75) = P(Z < 0.75) = 0.7734$.

La probabilità che su due studenti indipendenti almeno 1 sia un successo è la probabilità che non ci siano insuccessi $= 1 - P(2 \text{ insuccessi}) = 1 - q^2$. Siccome $q = 1 - 0.7734 = 0.2266$ la risposta è $1 - q^2 = 1 - (0.2266)^2 = 0.9487$.

DIFFICILE

Esercizio 6.38

Sappiamo che $X \sim N(\mu = 70, \sigma)$ ma la varianza non è nota. Quindi usiamo l'informazione

$$P(X < 85) = 0.9322$$

cioè standardizzando

$$P(Z < (85 - 70)/\sigma) = 0.9322$$

Poiché dalle Tavole della normale (Tavola 1 del Libro) si sa che $P(Z < 1.5) = 0.9322$ avremo l'equazione

$$1.5 = (85 - 70)/\sigma$$

che risolta ci dà

$$\sigma = 10.$$

La probabilità che $P(X > 80)$ è

$$P(Z > (80 - 70)/10) = P(Z > 1) = 1 - P(Z < 1) = 0.1587.$$

Ora, se poni $p = 0.1587$ come probabilità di 1 successo in 1 prova dove successo = lo studente supera 80 punti, l'esercizio chiede di trovare

$$\begin{aligned} P(1,2,3,4 \text{ successi su 4 prove indipendenti e identiche}) &= 1 - P(0 \text{ successi in 4 prove}) \\ &= 1 - \binom{4}{0} p^0 q^4 \\ &= 1 - 1 \cdot (1 - 0.1587)^4 = 0.4990394. \end{aligned}$$

La formula sopra deriva dal fatto che il numero di successi ha distribuzione Binomiale!!

È abbastanza DIFFICILE.

Esercizio 6.40

Il numero di successi è approssimabile con una normale con media $np = 1600 \cdot 0.4 = 640$ e deviazione standard $\sqrt{npq} = \sqrt{1600 \cdot 0.4 \cdot 0.6} = 19.6$.

a. Calcola
$$P(X > 690) \simeq P(Z > (690 - 640)/19.6) = P(Z > 2.55)$$
cioè $1 - P(Z < 2.55) = 0.0054$.

b. Calcola

$$P(X < 570) \simeq P(Z < (570 - 640)/19.6) = P(Z < -3.57) = 1 - 0.9888 = 0.0002.$$

c. Calcola

$$P(550 < X < 650) = P((550 - 640)/19.6 < Z < (650 - 640)/19.6) = P(-4.59 < Z < 0.51)$$

Ora questa probabilità è
$$P(Z < 0.51) - P(Z < -4.59) = 0.695 - 0 = 0.695.$$

d. Qual è il numero di successi che ha probabilità 0.9 di non essere superato?

$X =$ numero di successi $\approx N(640, \sigma^2 = 384)$ e quindi $\sigma = 19.6$.

Si chiede $P(X < x) = 0.09$. Cerca sulle tavole della $N(0, 1)$ il valore z tale che $P(Z < z) = 0.09$.

Questo valore non c'è perché è < 0.5. Allora occorre cercare il valore simmetrico rispetto allo 0 cioè $P(Z < -z) = 0.91$. Si trova $P(Z < 1.34) \simeq 0.91$, e quindi $P(Z < -1.34) = 0.09$.

Quindi ritrasformando per X si ha $x = 640 - (19.6)(1.34) = 613.74$. Quindi la risposta è 614 successi.

e. Qual è il numero di successi che ha probabilità 0.20 di essere superato?

Chiede: trova x tale che $P(X > x) = 0.20$. Si cerca allora sulle tavole della normale standard $P(Z < z) = 0.80$. Si trova $P(Z < 0.84) = 0.80$. Quindi ritrasformando per X si ha
$$x = 640 + (19.6)(0.84) = 656.46$$
cioè 656 successi.

Esercizio 6.42

d. Usa per la distribuzione di $\hat{P}$ una approssimazione normale con media 0.4 e $\sigma = \sqrt{pq/n} = 0.01224745$.

Poi hai che
$$P(\hat{P} < p) = P(Z < (p - 0.4)/0.01224745) = 0.2$$

Perciò poni $k = (p - 0.4)/0.01224745$ e poi osservi che $P(Z < k) = 0.2$ è equivalente a

$$P(Z < -k) = 0.8$$

Quindi dalle tavole trovi che $-k = 0.84$ (nota il **MENO**), ossia
$$-(p - 0.4)/0.01224745 = 0.84$$
e infine risolvi l'equazione per trovare p
$$p = 0.4 - 0.84 * 0.01224745 = 0.3897121$$

e. Farlo per esercizio perché è uguale al precedente.

Esercizio 6.48

La durata è $X \sim N(\mu = 35000, \sigma 4000)$. Sia

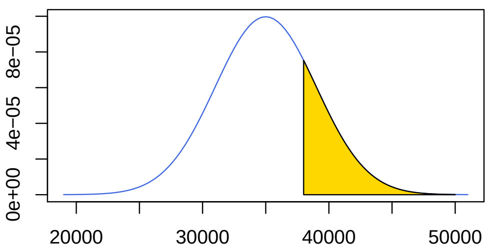
$$p = P(\bar{X} > 38000)$$

La probabilità che 1 pneumatico duri più di 38000 km si calcola

$$p = P(Z > (38000 - 35000)/4000) = P(Z > 0.75)$$

Quindi

$$p = 1 - P(Z < 0.75) = 0.22663$$



Quindi p è la probabilità di successo in 1 prova. In un campione di $n = 100$ prove indipendenti e identiche qual è la probabilità che ci siano più di 25 successi? Dobbiamo calcolare

$$P(X > 25)$$

dove X (il numero di successi) è binomiale con $n = 100$ e $p = 0.22663$. Il calcolo lo facciamo approssimando (visto che n è grande) X con una normale con media e varianza rispettivamente

$$\begin{aligned} \mu &= np = (100)(0.22663) = 22.663 \\ \sigma^2 &= npq = (100)(0.22663)(1 - 0.22663) = 17.52688 \end{aligned}$$

Quindi abbiamo

$$\begin{aligned} P(X > 25) &= P(Z > (25 - 22.663)/\sqrt{17.52688}) \\ &= P(Z > 0.56) \\ &= 1 - P(Z < 0.56) = 0.2877 \end{aligned}$$

E' DIFFICILE

Esercizio 6.67

È un esercizio facile. Se $T = X + Y + Z$ è il guadagno totale allora

$$E(T) = \mu_X + \mu_Y + \mu_Z = 20000 + 25000 + 15000 = 60000$$

e visto che sono X, Y e Z indipendenti

$$\text{var}(T) = \sigma_X^2 + \sigma_Y^2 + \sigma_Z^2 = 2000^2 + 5000^2 + 4000^2$$

Quindi la deviazione standard è

$$\sigma_T = \sqrt{2000^2 + 5000^2 + 4000^2} = 6708.2$$

Esercizio 6.73

Abbiamo che il numero di unità vendute è $X \sim N(\mu_X = 400, \sigma_X^2 = 900)$, e il numero di unità prodotte è $Y \sim N(\mu_Y = 400, \sigma_Y^2 = 1600)$.

Inoltre il prezzo di vendita unitario è 10 dollari. Quindi il ricavo è $10X$. Il costo unitario di produzione è 4 dollari, quindi il costo fisso di produzione è $4Y$.

La differenza tra i due è

$$W = 10X - 4Y$$

Siccome X e Y sono normali allora W è normale. Rispetto alle solite applicazioni qui però le due normali sono correlate. Questo caso non l'abbiamo mai visto a lezione.

Comunque possiamo calcolare la media e la varianza della variabile W usando le regole viste per le variabili doppie discrete. Perciò

$$E(W) = 10\mu_X - 4\mu_Y = 10(400) - 4(400) = 2400$$

$$\text{var}(W) = 100\sigma_X^2 + 16\sigma_Y^2 - 2(10)(4)\sigma_{XY}$$

Dobbiamo trovare la covarianza che è

$$\sigma_{XY} = \rho_{XY}\sigma_X\sigma_Y = 0.5(30)(40) = 600$$

Quindi

$$\text{var}(W) = 100(900) + 16(1600) - 2(10)(4)(600) = 67600$$

e dunque

$$\sigma_W = \sqrt{67600} = 260.$$

A questo punto la domanda finale chiede qual è la probabilità

$$P(W > 2000)?$$

Siccome $W \sim N(\mu_W = 2400, \sigma_W = 260)$ abbiamo

$$P(W > 2000) = P(Z > (2000 - 2400)/260) = P(Z > -1.54)$$

e questa è uguale a

$$P(Z < 1.54) = 0.9382$$

È DIFFICILE.

Esercizio 6.90

$$P(Z < \frac{50 - 70}{\sqrt{70}} = P(Z < -2.39) = 1 - F_Z(2.39) = 0.0084.$$

Capitolo 7. Campionamento distribuzioni campionarie stima

Esercizio 7.2

a. Binomiale con $n = 2, p = 0.5$ Funzione di probabilità

```
x        P(X=x)
0         0.25
1         0.50
2         0.25
```

b. Binomiale con $n = 4, p = 0.5$ Funzione di probabilità

x	P(X=x)
0	0.0625
1	0.2500
2	0.3750
3	0.2500
4	0.0625

c. Binomiale con $n = 10, p = 0.5$

Funzione di probabilità

x	P(X=x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977

Esercizio 7.5

a. La media è uguale alla media della popolazione cioè 100 e la varianza è uguale alla varianza della popolazione divisa per la numerosità campionaria cioè $81/25 = 3.24$. Quindi $\sigma_{\bar{X}} = \sqrt{3.24} = 1.8$.

b . Abbiamo

$$P(\bar{X} > 102) = P(Z > (102 - 100)/1.80) = P(Z > 1.11).$$

Dalle tavole si trova che il risultato è 0.1335.

c. Abbiamo

$$\begin{aligned} P(98 \leq \bar{X} \leq 101) &= P((98 - 100)/1.8 \leq Z \leq (101 - 100)/1.8) \\ &= P(-1.11 \leq Z \leq 0.56) \\ &= P(Z \leq 0.56) - [1 - P(Z \leq 1.11)] \\ &= 0.7123 - 0.1335 = 0.5788. \end{aligned}$$

d. Risulta

$$\begin{aligned} P(\bar{X} \leq 101.5) &= P[Z \leq (101.5 - 100)/1.8] \\ &= P(Z \leq 0.83) = 0.7967 \end{aligned}$$

Esercizio 7.12

Sia X la variabile *prezzo di vendita delle nuove abitazioni*; X ha distribuzione incognita, media e deviazione standard note: $X \sim N(\mu = 115000; \sigma = 25000)$.

Sia $n = 100$ l'ampiezza di un campione casuale estratto da X.

a. Essendo il campione sufficientemente ampio, si applica il teorema del limite centrale ottenendo

$$\bar{X} \sim N(115000, \sigma_{\bar{X}} = 25000\sqrt{100} = 2500)$$

La probabilità richiesta è

$$P(\bar{X} > 110000) = P(Z > (110000 - 115000)/2500) = P(Z > -2) = 0.9772$$

b. La probabilità

$$113000 < X < 117000) = P(-0.4 < Z < 0.4) = 0.6554 - (1 - 0.6554) = 0.3108$$

d. Essendo gli intervalli di stessa ampiezza e sfruttando la simmetria della distribuzione normale rispetto alla sua media 115000, $[114000; 116000]$ è l'intervallo in cui è più probabile trovare il prezzo medio campionario.

e. Quanto svolto nei punti precedenti è valido in quanto non è stato necessario attribuire alcuna distribuzione alla variabile X. Qualunque sia la distribuzione di X, è possibile applicare il teorema del limite centrale (perché la numerosità $n = 100$ del campione è sufficientemente grande) per il quale la media campionaria X ha con buona approssimazione distribuzione normale.

Esercizio 7.16

a. $ES = \sigma/\sqrt{100} = 40/10 = 4$.

b. Media campionaria $= \bar{X} \approx N(\mu, \sigma = 40/10 = 4)$ in quanto il campione è grande. Nota che: $Z = (\bar{X} - \mu)/4 \sim N(0, 1)$.

Quindi, chiede $P(\bar{X} > \mu + 5)$. Ossia standardizzando,

$$P(Z > (\mu + 5 - \mu)/4) = P(Z > 5/4) = 1 - P(Z \leq 1.25) = 1 - 0.8944 = 0.1056.$$

c. $P(\bar{X} - \mu > 4) = P(4Z > 4) = 1 - P(Z \leq 1) = 1 - 0.8413$.

d. Questa volta chiede che la distanza tra la media campionaria $\bar{X}$ e μ sia maggiore di 3. Si usa perciò il valore assoluto:

$$\begin{aligned} P(|\bar{X} - \mu| > 3) &= P(4|Z| > 3) \\ &= P(Z \leq -0.75) + P(Z > 0.75) \\ &= 2(1 - 0.7734) = 0.4532. \end{aligned}$$

Esercizio 7.18

X = livello di impurità $\sim N(\mu, \sigma = 1.6)$; Campione casuale $n = 100$. $\bar{X} \sim N(\mu, \sigma/\sqrt{n} = 0.16)$.

a. Sia k un incremento della media della popolazione. Si chiede: trova k tale che

$$P(\bar{X} > \mu + k) = 0.05$$

cioè

$$P(\bar{X} - \mu > k) = 0.05$$

Divido ambo i membri della disuguaglianza per $\sigma/\sqrt{n} = 0.16$ e ottengo

$$P(Z > k/0.16) = 0.05$$

Perciò poiche nella normale standard $P(Z > h) = 0.05$ si ha per $h = 1.645$ dovremo avere qui $k/0.16 = 1.645$ e cioè $k = 0.16 \cdot 1.645 = 0.2632$.

b. Chiede $P(\bar{X} < \mu - k) = 0.10$. Quindi $P(Z < -k/0.16) = 0.10$.

Nella normale standard $P(Z < h) = 0.10$ si ha per $h = -1.282$.

ATTENZIONE: guarda le tavole della t di Student ultima riga: $P(Z > 1.282) = 0.10$.

Quindi qui occorre che $-k/0.16 = -1.282$, cioè $k = 0.16 \cdot 1.282 = 0.205$. Il libro dà come soluzione -0.205, ma si sbaglia. La media campionaria deve essere inferiore di 0.205 alla media della popolazione non di MENO 0.205.

c. Chiede $P(|\bar{X} - \mu| > k) = 0.15$. Quando dice "differire" vuol dire sia per valori superiori che inferiori e quindi bisogna usare il valore assoluto.

Quindi $P(|Z| > k/0.16) = 0.15$. Poiché nella Normale standard $P(|Z| > k) = 0.15$ si ha per $k = 1.44$.

ATTENZIONE: guarda le tavole della Normale.

Per $z = 1.44$ $P(Z < 1.44) = 0.925$. Quindi a destra di 1.44 ci sta $1 - 0.925 = 0.075 = 7.5\%$. Analogamente a sinistra di -1.44 ci sta 7.5% e quindi complessivamente nelle due code ci sta il 15%. Guarda la figura: le due code sono delimitate da -1.44 e 1.44.

Pertanto nel nostro caso $k/0.16 = 1.44$ e cioè $k = 0.16 \cdot 1.44 = 0.2304$.

Esercizio 7.20

Ore $= X \sim N(\mu, \sigma = 8.4)$. Dimensione campionaria n incognita. La distribuzione campionaria della media è

$$\bar{X} \sim N(\mu, 8.4/\sqrt{n}).$$

a. Chiede di trovare n tale che $P(|\bar{X} - \mu| > 2) \leq 0.05$.

Scriviamo la condizione in un altro modo:

$$P(\bar{X} > \mu + 2 \text{ oppure } \bar{X} < \mu - 2) \leq 0.05$$

Quindi possiamo scrivere anche

$$P(\bar{X} > \mu + 2) + P(\bar{X} < \mu - 2) \leq 0.05$$

Poi si nota anche che

$$P(\bar{X} > \mu + 2) = P(\bar{X} < \mu - 2)$$

Quindi la condizione che è data all'inizio si può sostituire con

$$2P(\bar{X} > \mu + 2) \leq 0.05.$$

Calcolo la probabilità a primo membro standardizzando rispetto a media e varianza di $\bar{X}$.

$$P(\bar{X} > \mu + 2) = P(Z > \frac{\mu + 2 - \mu}{8.4/\sqrt{n}}) = P(Z > \sqrt{n}\frac{2}{8.4})$$

Quindi la condizione che deve essere rispettata è

$$2P(Z > \sqrt{n}\frac{2}{8.4}) \leq 0.05$$

ossia

$$P(Z > \sqrt{n}\frac{2}{8.4}) \leq 0.025$$

Pongo $h = \sqrt{n}\frac{2}{8.4}$ e cerco di risolvere l'equazione

$$P(Z > h) = 0.25 \text{ ossia } P(Z \leq h) = 0.975$$

Le tavole della normale ci dicono che deve essere (h = 1.96). Perciò

$$h = \sqrt{n}\frac{2}{8.4} = 1.96$$

Si risolve rispetto all'incognita n ottenendo

$$\sqrt{n} = \frac{8.4 \cdot 1.96}{2} = 8.232$$

da cui $n = 67.77$. Perciò basta che $n = 68$ affinché la richiesta sia soddisfatta. All'aumentare di n poiché $\sigma/\sqrt{n}$ decresce e la probabilità nelle code diminuisce, si otterrà di sicuro che la probabilità richiesta sia ≤ 0.05.

b. Se si cambia 0.05 in 0.10 ovviamente invece di 1.96 basterebbe un valore più piccolo e quindi n è minore.

c. Se si cambia 0.05 in 0.01 invece di 1.96 ci vorrebbe un valore più grande e quindi un n maggiore.

Esercizio 7.38

SE $= \sqrt{pq/100}$. Il massimo errore si ha per $p = 0.5$ cioè

$$\max \mathrm{SE} = 0.5/10 = 0.05.$$

Esercizio 7.39

Chiede di trovare n tale che $P(|\text{Proporzione} - p| > 0.03) = 0.05$. Quindi siccome Proporzione $\approx N(p, \sigma = 0.5/\sqrt{n})$, allora $Z = \sqrt{n}(\text{Proporzione} - p)/0.5 \sim N(0,1)$ ossia

$$\text{Proporzione} - p = Z/(2\sqrt{n}).$$

Allora

$$P(|\text{Proporzione} - p| > 0.03) = P(|Z| > 0.03 \cdot [2 \cdot \sqrt{n}]) = 0.05.$$

Siccome sappiamo che $P(|Z| > 1.96) = 0.05$ occorre che

$$0.03 \cdot [2 \cdot \sqrt{n}] = 1.96$$

e quindi $\sqrt{n} = 1.96/0.06 = 32.67$ e dunque $n > 32.67^2 = 1067.1$.

Esercizio 7.44

a. Abbiamo

$$\hat{p} = \frac{211}{528} = 0.3996$$

Quindi

$$\sigma_{\hat{p}} = \sqrt{\frac{(0.3996)(0.6.004)}{120}} = 0.03934.$$

b. Risulta che la probabilità cercata è

$$P(Z < (0.33 - 0.3996)/0.03934) = P(Z < -1.77) = 0.0384$$

c. La probabilità è dopo aver standardizzato

$$P((0.5 - 0.3996)/0.03934 < Z < (0.6 - 0.3996)/0.03934) = P(2.55 < Z < 5.09)$$

Quindi siccome $P(Z < 5.09) = 1$ risulta

$$P(Z < 5.09) - P(Z < 2.55) = 1 - 1 - 0.9946 = 0.0054.$$

Esercizio 7.46

ESERCIZIO DIFFICILE

Sia Y l'aumento di stipendio. Allora $Y \sim N(12, \sigma = 3.6)$. Estraggo a caso un dirigente e dico che ottengo un successo se l'aumento è (Y < 10%). La probabilità di successo è

$$p = P(Y < 10) = P(Z < (10 - 12)/3.6) = P(Z < -0.56).$$

Quindi

$$p = 1 - P(Z < 0.56) = 1 - 0.7123 = 0.2877.$$

Poi considero un campione casuale (con ripetizione) di dirigenti dimensione $n = 81$ e guardo la proporzione di successi $\hat{P}$ di successi nel campione.

L'esercizio chiede di calcolare la probabilità $P(\hat{P} > 0.5)$.

Ora $\hat{P}$ ha distribuzione binomiale con media $p = 0.2877$ e deviazione standard

$$\sigma_P = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.2877)(0.7123)}{81}} = 0.0503.$$

Quindi usando l'approssimazione normale alla Binomiale

$$P(\hat{P} > 0.5) \approx P(X > 0.5)$$

dove $X \sim N(p, \sigma_P)$. Pertanto

$$P(X > 0.5) = P(Z > (0.5 - 0.2877)/0.0503) = P(Z > 4.22) = 0.$$

Esercizio 7.66

- Popolazione: $X \sim N(420, \sigma = 100)$
- Campione casuale: $n = 25$
- $\bar{X} \sim N(420, \sigma = 100/\sqrt{25} = 20)$

a. $P(\bar{X} > 450) = 1 - P(Z < (450 - 420)/20) = 1 - P(Z < 1.5) = 0.0668.$

b. Risulta

$$\begin{aligned} P(400 < \bar{X} < 450) &= P(\bar{X} < 450) - P(\bar{X} < 400) \\ &= (1 - 0.0668) - P(Z < (400 - 420)/20) \\ &= 0.93319 - P(Z < -1) = 0.93319 - (1 - P(Z < 1)) \\ &= 0.93319 - 0.1587 = 0.7745. \end{aligned}$$

c. Significato: qual è il valore c tale che $P(\bar{X} > c) = 0.10$? È uguale a $P(\bar{X} \leq c) = 0.9$. Per la normale standard il valore z tale che $P(Z > z) = 0.10$ è $z = 1.28$ (guardate sulle tavole). Ma $\bar{X}$ ha distribuzione $N(420, \sigma = 100/\sqrt{25} = 20)$ e quindi dovete trasformarlo con

$$c = 420 + 20 \cdot 1.28 = 445.6.$$

d. Significato: Qual è il valore z tale che $P(\bar{X} < 0.10)$? Siccome per la normale standard il valore z tale che $P(Z \leq z) = 0.10$ è $z = -1.28$ il valore cercato è $c = 420 + 20 \cdot (-1.28) = 394.4$.

e. Significato: qual è il valore k tale che $P(\text{Deviazione stand.} > k) = 0.05$ ossia $P(S^2 > k^2) = 0.05$?

Per risolvere questo esercizio bisogna sapere la distribuzione campionaria (chi-quadro) della varianza, che però non è in programma.

f. Stesso problema dell'esercizio e.

g. Sarebbe stata minore perché la distribuzione di $\bar{X}$ avrebbe avuto la stessa media, ma varianza più piccola e quindi le code si sarebbero ritirate.

Esercizio 7.72

a. P = proporzione che accetta di lavorare in un campione di $n = 60$ quando nella popolazione $p = 0.8$. È Binomiale relativa : $P \sim \mathit{Binomiale}(p = 0.8, n = 60)/60$.

Poiché $npq = 60 \cdot 0.8 \cdot 0.2 = 9.6 > 9$ si può usare l'approssimazione normale.

$$P \approx N(\mu = p = 0.8, \sigma = \sqrt{(0.8 \cdot 0.2)/60} = 0.05163978).$$

Perciò

$$P(P < 0.7) = P(Z < (0.7 - 0.8)/0.05163978) = P(Z < -1.94) = 0.0262.$$

b. Se il campione fosse di 6 soltanto occorre usare la Binomiale perché l'approssimazione normale non va bene.

$$P(P < 0.7) = P(\text{successi}/6 < 0.7) = P(\text{successi } < 4.2) = 1 - P(\text{successi} = 5 \text{ o } 6).$$

Quindi

$$P(\text{successi} = 5) = \binom{6}{5} p^5 q^1 = 6 \cdot 0.8^5 \cdot 0.2 = 0.393216$$

$$P(\text{successi} = 6) = p^6 = 0.8^6 = 0.262144$$

e infine

$$P(P < 0.7) = 1 - 0.393216 - 0.262144 = 0.34464.$$

c. In un campione di $n = 6$, $\bar{X}$ = stipendio medio $\sim N(29000, 4000/\sqrt{6} = 1632.993 = \sigma(\bar{X}))$

$$P(\bar{X} > 30000) = P(Z > (30000 - 29000)/1632.993) = P(Z > 0.61) = 0.2709.$$

d. Risulta

$$\begin{aligned} P(\text{Accettato e Stipendio} > 30000) &= P(\text{Accettato})P(\text{ Stipendio} > 30000 \mid \text{Accettato}) \\ &= 0.8 \cdot P[Z > (30000 - 29000)/4000] = 0.8 \cdot P(Z > 0.25) \\ &= 0.8 \cdot 0.4013 = 0.3210. \end{aligned}$$

Capitolo 8. Stima e intervalli di confidenza

Esercizio 8.2

a. L'unico modo che conoscete per controllare se i dati sono un campione da una normale è fare un box-plot. Il box-plot segnala una lieve asimmetria che naturalmente può sparire in un campione più grande.

b. Uno stimatore non distorto (cioè corretto) ed efficiente della media μ della popolazione è la media campionaria $\bar{X}$. Quindi

$$\bar{x} = 101.375.$$

c. Uno stimatore non distorto della varianza σ^2 della popolazione è la varianza campionaria corretta

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = 201.6964.$$

Perciò l'errore standard dello stimatore della media è

$$ES = \frac{s}{\sqrt{n}} = \sqrt{201.6964/8}.$$

d. Sia p la proporzione nella popolazione di case in vendita a meno di \$92500. Questa si può stimare con la proporzione analoga valutata nel campione. Lo stimatore in questione è non distorto. Si calcola con

$$\hat{P} = \frac{\#\text{case a meno di } 92500}{n} = 3/8.$$

Infatti nel campione ci sono 3 pressi su 8 inferiori a \$92500.

Esercizio 8.9

Il fattore di affidabilità è $z_{\alpha/2}$. Per trovare il valore dalle tavole si cerca il valore k tale che

$$P(Z < k) = 1 - \alpha/2$$

a. $1 - \alpha = 0.96$ quindi $\alpha/2 = 0.02$. Quindi $z_{\alpha/2} = 2.05$. Infatti

$$P(Z < 2.05) \simeq 0.98$$

b. $\alpha/2 = 0.06$. Quindi $z_{\alpha/2} = 1.56$.

c. $\alpha/2 = 0.075$. Quindi $z_{\alpha/2} = 1.44$.

d. $\alpha/2 = 0.035$. Quindi $z_{\alpha/2} = 1.81$.

e. $\alpha/2 = 0.07$. Quindi $z_{\alpha/2} = 1.48$.

Esercizio 8.14

a. $\alpha/2 = 0.04$. Quindi $z_{\alpha/2} = 1.75$.

b. $ES = 6/\sqrt{90} = 0.63246$.

c. Ampiezza $= 2ME = 2(1.75)ES = 2.2136.$.

Esercizio 8.22

a. n= 6 Livello = 1 - alfa = 95%. Sulle tavole trovate $t = 2.271$ Quindi $ME = t \cdot s/\sqrt{n} = 2.571 \cdot 40/\sqrt{6} = 41.98$. L'ampiezza è $2 * ME = 2 \cdot 41.98 = 83.96$.

b. $n = 22, s = 20$, Livello = 99%, t con 21 gdl = 2.831. Quindi $2ME = 2 \cdot 20/\sqrt{22} \cdot 2.831 = 24.1428$.

c. $n = 25$, $s = 50$, Livello = 90% t con 24 gdl = 1.714.

Quindi $2ME = 2 \cdot 50/5 \cdot 1.714 = 34.2$.

Esercizio 8.38

Il campione ha numerosità $n = 198$. Si vuole stimare la proporzione p di quelli che giudicano non etico il comportamento. La proporzione campionaria è $\hat{p} = 98/198 = 0.495$. L'intervallo di confidenza è (ci dice) uguale a $(0.445, 0.545)$. Quindi il margine di errore è

$$ME = (0.545 - 0.445)/2 = 0.05 = z_{\alpha/2} ES$$

dove

$$ES = \sqrt{\hat{p}(1-\hat{p})/n} = 0.03553164$$

è l'errore standard. Perciò

$$z_{\alpha/2} = ME/ES = 0.05/0.03553164 = 1.41$$

Quindi dalla Tavola 1, a p.521

$$\alpha/2 = 1 - F(1.41) = 1 - 0.9207 = 0.0793$$

e inoltre

$$\alpha = 2 \times 0.0793 = 0.1586$$

Perciò il livello di confidenza è

$$1 - \alpha = 1 - 0.1586 = 0.8414.$$

cioè l'84.14%.

Esercizio 8.47

Qui ci sono i dati grezzi e il campione è molto piccolo. La popolazione è incognita (non è detto che sia normale)

a. Le stime più efficienti (se la popolazione è normale) sono la media e la varianza campionaria corretta:

$$\bar{x} = 3.375, \quad s^2 = 0.4993.$$

$

b. Sia Y la variabile aleatoria che vale 1 se il livello di impurità di un sacco è superiore a 3.75% e 0 altrimenti. Questa variabile è Bernoulli con probabilità di successo p incognita. p è la proporzione nella popolazione di sacchi con un livello di impurità superiore al 3.75%. Il campione estratto da X genera il seguente campione da Y:

$$(0, 1, 0, 0, 0, 0, 0, 1, 1)$$

in quanto solamente la seconda, la settima e l'ottava osservazione sono superiori a 3.75%. Perciò la stima di p è

$$\hat{p} = 3/8.$$

Esercizio 8.48

Media campionaria = 3.59. L'errore standard è $1.045/\sqrt{457} = 0.04888304$.

Quindi l'IC è

$$(3.59 - t \cdot 0.04888304, 3.59 + t \cdot 0.04888304) = (3.49, 3.69)$$

Questo vuol dire che $0.1 = t \cdot 0.04888304$ e quindi $t = 0.1/0.04888304 = 2.045699 = 2.05$.

Siccome il campione ha dimensione $n = 457$, la t di Student con 456 gdl è praticamente uguale alla normale $N(0, 1)$.

Sulle tavole della normale si vede che 2.05 corrisponde ad una probabilità cumulata 0.9798, cioè $P(Z < 2.05) = 0.9798$.

Quindi $\alpha/2 = 1 - 0.9798 = 0.0202$. Il livello di confidenza $1 - \alpha$ è dunque $1 - 2 \cdot 0.0202 = 0.9696 = 95.96\%$.

Esercizio 8.49

$n = 174$, $\bar{x} = 6.06$, $s = 1.43$. Quindi

$$6.16 - 6.06 = 0.1 = z_{\alpha/2}(1.43/\sqrt{174})$$

e

$$z_{\alpha/2} = 0.922$$

Quindi

$$\alpha = 2[1 - P(Z < 0.92)] = 0.3576$$

e

$$100(1 - 0.3576)\% = 64.24\%$$

Capitolo 10. Test delle ipotesi

Esercizio 10.11

X = peso della confezione in once. Si sa che $X \sim N(\mu, \sigma = 0.4)$. - Il sistema di ipotesi è $H_0 : \mu \geq 16; H_1 : \mu < 16$. - Campione casuale con $n = 16$. - Peso medio $\bar{x} = 15.84$. - Errore standard $ES = \sigma/\sqrt{n} = 0.4/4 = 0.1$. - Statistica test: $Z = \frac{\bar{X}-16}{0.1}$ - Livello = 0.1 - Regione critica: $Z < -z_{0.1} = -1.28$.

Quindi siccome $Z = \frac{15.84-16}{0.1} = -1.6$ si rifiuta l'ipotesi al livello del 10%.

Esercizio 10.17

Il giudizio è $X \sim N(\mu, \sigma)$. L'ipotesi che la popolazione sia normale è alquanto discutibile.

- Il sistema di ipotesi è $H_0 : \mu = 4; H_1 : \mu \neq 4$.
- Campione casuale con dimensione campionaria $n = 1562$ quindi molto alta.
- Giudizio medio campionario $\bar{x} = 4.27$.
- Deviazione standard campionaria $s = 1.32$.
- Errore standard $ES = s/\sqrt{n} = 1.32/\sqrt{1562} = 0.033399$
- Statistica test $= T = \frac{\bar{X}-4}{ES}$
- Livello del test $= 1\%$
- Gradi di libertà $= g = 1562 - 1 = 1561$. Sono elevati. La distribuzione campionaria di T è comunque normale standard sotto ipotesi nulla in virtù del teorema centrale del limite.
- Regione critica (T > t_{0.01/2,g} oppure T < -t_{0.01/2,g}).

Decisione: siccome $T = 8.08$ e $t_{0.01/2,g} = 2.576$ si rifiuta l'ipotesi. Si rifiuterebbe anche a un livello più piccolo di α.

Esercizio 10.26

- Incremento di vendite $= X \sim N(\mu, \sigma^2)$.
- Sistema di ipotesi: $H_0 : \mu \geq 50; H_0 : \mu < 50$ Il testo dice "si realizzerà un incremento medio di 50 casse" perciò si potrebbe anche usare $H_0 : \mu = 50$ ma in tal caso l'affermazione verrebbe falsificata se l'incremento medio fosse invece che di 50 casse, di 60 o più casse e ciò ovviamente non sembra molto ragionevole. Per cui ho optato per un ipotesi nulla $H_0 : \mu \geq 50$.
- Campione casuale di dimensione $n = 20$.
- Media campionaria $\bar{x} = 41.3$.
- Deviazione standard campionaria $s = 12.2$.
- Errore standard $ES = s/\sqrt{n} = 12.2/\sqrt{20} = 2.728$.
- Statistica test: $T = \frac{\bar{X}-50}{ES}$ distribuita sotto ipotesi nulla come una t di Student con 19 gradi di libertà.
- Livello del test (5%)
- Gradi di libertà =19
- Valore critico per un test unilaterale sinistro $-t_{0.05,19} = -1.729$
- Decisione: Poiché $t = \frac{41.3-50}{2.728} = -3.189$ si rifiuta l'ipotesi al livello del 5%.

Esercizio 10.35

- L'ipotesi nulla è $H_0 : p \geq 0.75$ contro $H_1 : p < 0.75$ dove p è la proporzione di insegnanti che condividono l'affermazione.
- La stima di $\hat{p} = 118/172 = 0.686$.
- L'errore standard sotto ipotesi nulla è

$$ES = \sqrt{\frac{0.75 \cdot 0.25}{172}}$$

- La statistica test è

$$z = \frac{0.686 - 0.75}{ES} = -1.94$$

- Il p-value è
$$P(Z < -1.94) = P(Z > 1.94) = 1 - P(Z < 1.94) = 0.0262$$
Quindi il test è significativo ma non altamente significativo. In particolare il test è significativo al livello del 5%.

Esercizio 10.40

a. L'ipotesi $H_0 : \mu \leq 3$ contro $H_1 : \mu > 4$ si rifiuta al livello del 5% se
$$\frac{\bar{X} - 3}{0.4/\sqrt{64}} > 1.645$$
cioè se $\bar{X} > 3.082$. Poiché nel campione $\bar{X} = 3.07$ la decisione è non rifiutare l'ipotesi nulla.

b. La probabilità di rifiutare H_0 quando è falsa e $mu = 3.1$ con la regola precedente è
$$P(\bar{X} > 3.082)$$
però con $\bar{X} \sim N(3.1, \sigma_{\bar{X}} = 0.4/\sqrt{64})$. Quindi
$$P(Z > \frac{3.082 - 3.1}{0.4/\sqrt{64}}) = P(Z > -0.36) = P(Z < 0.36) = 0.6406.$$

Esercizio 10.42

La popolazione è dicotomica. Cioè $X = 1$ se il cliente preferisce prodotti non di marca a prezzo infereiore. Quindi $X \sim Bernoulli(p)$.

- Il sistema di ipotesi da verificare è $H_0 : p \geq 0.5; H_1 : p < 0.5$.
- Campione casuale di dimensione $n = 802$.
- Proporzione campionaria $= \hat{p} = 378/802 = 0.4713217$.
- Errore standard sotto $p = 0.5$: $ES_0 = \sqrt{p_0(1 - p_0)/n} = \sqrt{(0.5)(0.5)/802} = 0.01765561$.
- Statistica test $(\hat{P} - 0.5)/ES_0$ distribuita sotto $p = 0.5$ approssimativamente come una normale standard per il teorema centrale del limite.
- Livello del test = 10%
- Valore critico per un test unilaterale sinistro (-z_{0.1} = -1.28).

Quindi si rifiuta al livello del 10% se
$$\frac{\hat{P} - 0.5}{0.01765561} < -1.28 \text{ ossia se } \hat{P} < 0.477.$$
Perciò in questo caso si rifiuta H_0 al livello del 10%.

La potenza del test chiesta è la probabilità di rifiutare H_0 quando $p = 0.45$. Quindi è
$$P(\hat{P} < 0.477) \text{ se } p = 0.45$$
Notate che se $p = 0.45$ allora $\hat{P}$ è Binomiale con media $p = 0.45$ e deviazione standard
$$\sqrt{pq/n} = \sqrt{(0.45)(0.55)/802} = 0.01756711$$
e inoltre siccome la numerosità campionaria è elevata si può approssimare con una normale. Quindi la probabilità di rifiutare è
$$P(\hat{P} < 0.477) \approx P(Z < (0.477 - 0.45)/0.01756711) = P(Z < 1.537) = 0.9378.$$
La seconda parte È DIFFICILE.

Esercizio 10.45

a.
$$\alpha = P(\frac{30.8-32}{3/\sqrt{36}}) = P(Z < -2.4) = 0.0082$$

b.
$$\alpha = P(\frac{30.8-32}{3/\sqrt{9}}) = P(Z < -1.2) = 0.1151$$

c.
$$\beta = P(\frac{30.8-31}{3/\sqrt{36}}) = P(Z > -0.4) = 0.6554$$

Esercizio 10.52

Sia X la potenza giornaliera. Dunque $X \sim N(\mu, \sigma = 120)$. - Il sistema di ipotesi è

$$H_0 : \mu \geq 800, \quad H_1 : \mu < 800$$

- Il campione casuale ha numerosità $n = 100$. - La regione di rifiuto è definita da tutti i campioni cha hanno una media campionaria $\bar{X} < 776$.

a. La probabilità di errore del primo tipo è la probabilità di rifiutare se H_0 è vera cioè

$$P(I) = P(\bar{X} < 776;\ \text{quando } \mu = 800)$$

Quindi sotto ipotesi nulla $\bar{X} \sim N(\mu = \mu_0 = 800, \sigma_{\bar{X}} = \sigma/\sqrt{100} = 12)$.

Allora
$$\alpha = P(I) = P(Z < (776-800)/12) = P(Z < -2) = 1 - P(Z < 2) = 0.0228.$$

b. La probabilità di errore del secondo tipo se la vera media fosse $\mu = 740$ è la probabilità di accettare H_0 quando H_0 è falsa e $\mu = 740$ cioè

$$\beta = P(II) = P(\bar{X} \geq 776;\ \text{quando } \bar{X} \sim N(740, \sigma_{\bar{X}} = 12))$$

Quindi
$$P(II) = P(Z \geq (776-740)/12) = P(Z \geq 3) = 1 - P(Z < 3) = 0.0014.$$

c. Se fosse $n = 200$

d. α sarebbe minore di quello della risposta a. Avremmo infatti $\alpha = P(Z < -2.83)$.

ii. β sarebbe minore di quello calcolato per la risposta b. Avremmo infatti $\beta = P(Z \geq 4.24)$.

iii. Se fosse $n = 100$ e avessimo una regione critica $\bar{X} < 756$

iv. α sarebbe minore di quello calcolato per la risposta a. Avremmo $\alpha = P(Z < -3.67)$.

v. β sarebbe maggiore di quello calcolato per la risposta b.
Avremmo $\beta = P(Z \geq 1.33)$.

Esercizio 10.53

a. $H_0 : p \geq 0.25, \quad H_1 : p < 0.25$. Regione critica al 5%: Rifiuta se $z < -1.645$.

$$z = \frac{0.215 - 0.25}{\sqrt{0.25(0.75)/545}} = -1.9$$

quindi rifiuta al livello 5% .

b. H_0 si rifiuta se $z < -1.645$ ossia se $\hat{p} < .2195$.

i) Potenza

$$1P(Z > \frac{0.2195 - 0.2}{\sqrt{0.2(0.8)/545}}) = 1P(Z > 1.14) = 0.8729$$

ii) Potenza

$$1P(Z > \frac{0.2195 - 0.25}{\sqrt{0.25(0.75)/545}}) = 1P(Z > -1.64) = 0.0505$$

iii) Potenza

$$1P(Z > \frac{0.2195 - 0.3}{\sqrt{0.3(0.7)/545}}) = 1P(Z > -4.1) = 0.0000$$