Esercizi misti (2)

Esercizio 1

Un collettivo di 200 donne è stato classificato secondo lo stato civile e l'età come segue:

	Stato civile	
Età	Nubile	Coniugata
Fino a 25 anni	40	15
Più di 25 anni	45	100

Sia **A** l'evento "*avere età fino a 25 anni*" e **B** l'evento "*essere coniugata*"

a) calcolare P(A)
b) Calcolare P(B)
c) Calcolare $P(A \cup B)$
d) Calcolare $P(B|A)$
e) Calcolare $P(A \cap B)$
f) A e B sono eventi indipendenti?

Soluzioni

	Stato civile		**Totale**
Età	Nubile	Coniugata	
Fino a 25 anni	40	15	55
Più di 25 anni	45	100	145
Totale	85	115	200

a) p(A) = 55/200 = 0.275

b) p(B) = 115/200 = 0.575

c) $p(A \cup B) = p(A)+p(B)- p(A \cap B)= 0.275+0.575-15/200=0.775$

d) $p(B|A) = p(B \cap A)/p(A)= (15/200)/(55/200)=15/55=0.273$

e) $p(A \cap B) = 15/200 = 0.075$

f) A e B sono eventi indipendenti?

Due eventi si dicono indipendenti se $p(A|B)=p(A)$ o $p(B|A)=p(B)$ o $p(A \cap B)=p(A)*p(B)$.

Dai calcoli già eseguiti risulta che $p(B|A)=0.273 \neq p(B)=0.575$ e che $p(A \cap B)=0.075 \neq p(A)*p(B)=0.275*0.575=0.158$

Esercizio 2

Per un paziente con certi sintomi si considerino i seguenti eventi:

A:={ha l'influenza} B:={ha la polmonite} C:={ha la febbre a 40}

Sapendo che:

$A \cap B=\phi$ $A \cup B=\mathbf{I}$ $P(A)=0.7$ $P(C|A)=0.3$ $P(C|B)=0.8$

a) Qual è la probabilità che il paziente abbia la polmonite? (essendo **I** l'evento certo)

b) Qual è la probabilità che abbia l'influenza se ha febbre a 40?

Soluzioni

a) Essendo **I** l'evento certo, risulta che: p(B)=1-p(A)=1-0.7=0.3

b) $p(A|C)=p(A \cap C)/p(C)$

Allora **p(C)**$=p(C \cap A)+p(C \cap B)=p(C|A)*p(A)+p(C|B)*p(B)=$

$=0.3*0.7+0.8*0.3=0.21+0.24=0.45$

$\mathbf{p(A \cap C)}=p(C|A)*p(A)=0.3*0.7=0.21$

Quindi $p(A|C)=p(A \cap C)/p(C)=0.21/0.45=0.47$

Esercizio 3

Uno studente al primo anno di università vuole conoscere le sue possibilità di laurearsi entro 4 anni. Gli vengono fornite le seguenti informazioni:

1) Il 15% degli iscritti si laurea entro 4 anni
2) Su 10 laureati entro 4 anni, 6 hanno riportato il massimo dei voti all'esame di diploma di scuola media superiore
3) Su 100 laureati con tempi superiori a 4 anni, 10 hanno riportato il massimo dei voti all'esame di diploma di scuola media superiore

Sapendo che lo studente in questione ha riportato il massimo dei voti all'esame di diploma di scuola media superiore, qual è la probabilità che si laurei entro 4 anni?

Soluzione

Eventi:

A={laurea entro 4 anni}

B={voto massimo}

Sappiamo che:

$p(A)=0.15$

$p(B|A)=6/10=0.6$

$p(B|\bar{A})=10/100=0.1$

Devo trovare $\mathbf{p(A|B)=p(A \cap B)/p(B)}$

$\mathbf{p(A \cap B)}=p(B|A)*p(A)=0.6*0.15=0.09$

$\mathbf{P(B)}=p(B \cap \bar{A})+p(B \cap A)=p(B|\bar{A})*p(\bar{A})+p(B|A)*p(A)=$

$=0.1*0.85+0.6*0.15=0.085+0.09=0.175$

Dunque $\mathbf{p(A|B)=p(A \cap B)/p(B)}=0.09/0.175=0.514$

Esercizio 4

Il tempo di percorrenza di un tratto autostradale è descritto da una variabile casuale con la seguente distribuzione di probabilità:

Tempo (minuti)	**Probabilità**
[15, 20]	0.15
(20, 23]	0.25
(23, 27]	0.40
(27, 31]	0.20

a) Calcolare la probabilità di percorrere il tratto autostradale in non più di 23 minuti
b) Calcolare la probabilità di percorrere il tratto autostradale in un tempo T tale che $20<T<27$

Soluzioni

a) probabilità di percorrere il tratto autostradale in non più di 23 minuti. Sappiamo che:

p(A)=0.15

p(B)=0.25

p(C)=0.40

p(D)=0.20

$p(A \cup B)=0.15+0.25=0.40$ dato che (A e B) sono eventi mutuamente esclusivi

b) probabilità di percorrere il tratto autostradale in un tempo T tale che 20<T<27

$p(B \cup C)=0.25+0.40=0.65$

Esercizio 5

La quantità P in grammi di farina erogati in ogni confezione da una macchina si distribuisce normalmente con media 500g e σ=10g.

a) Calcolare la probabilità che vengano erogati meno di 485g.

b) Calcolare la probabilità che la quantità erogata sia compresa tra 490g e 512g.

c) Stabilire quel peso P0 per il quale la probabilità che la macchina eroghi una quantità di farina maggiore di P0 è pari a 0.14

Soluzioni

a) Sappiamo che **P**~N(500,100)

$p(\mathbf{P}<485)=p[(\mathbf{P}-\mu)/\sigma<(485-\mu)/\sigma]=p[z<(485-\mu)/\sigma]=$

$=p[z<(485-500)/10]=p(z<-1.5)=1-0.933=0.067$

Cioè il 6.7% di probabilità che vengano erogati meno di 485g

b) $p(490<\mathbf{P}<512)=p[(490-\mu)/\sigma<(\mathbf{P}-\mu)/\sigma<(512-\mu)/\sigma]=$

$p[(490-500)/10<z<(512-500)/10]=$

$=p[-1<z<1.2]=1-p(z>1)-p(z>1.2)=1-0.15866-0.11507=0.727$

La probabilità è del 72.7%

c) Sappiamo che $p(\mathbf{P}>P_0)=0.14$, dalle tavole ricavo che $z_0=1.08$, da cui:

$z_0=(P_0-\mu)/\sigma \Longrightarrow 1.08=(P_0-500)/10$, dunque $P_0=1.08*10+500=510.8g$

La probabilità che la macchina eroghi una quantità di farina maggiore di 510.8g è pari al 14%

Esercizio 6

Per vari motivi gli esiti delle malattie sono frequentemente auto-riferiti senza una verifica epidemiologica. In uno studio di Parikh-Patel et al., i ricercatori hanno voluto osservare la relazione che c'è tra i casi di tumore auto-riferito e i casi reali presenti. Hanno utilizzato i dati di cancro auto-riferito del California Teachers Study e per quelli dei casi presenti di cancro hanno utilizzato il registro tumori. La seguente tabella mostra i dati trovati.

	Tumore registrato		
Tumore autocertificato	Sì (B)	No (NB)	Totale
Sì (A)	2991	2244	5235
No (NA)	112	115849	115961
Totale	3103	118093	121196

Quesiti

Dati gli eventi:

A=tumore auto riferito nello studio California Teachers

B=tumore confermato dal Registro Tumori della California

a) Qual è la probabilità che si verifichi A?
b) Qual è la probabilità che si verifichi B?
c) Trovare la $P(A \cap B)$
d) Trovare la $P(A|B)$
e) Trovare la $P(B|A)$
f) Calcolare la Se nell'utilizzare le informazioni autoriferite di tumore come predittore del fatto che poi il tumore sia presente nel registro californiano
g) Calcolare la Sp nell'utilizzare le informazioni autoriferite di tumore come predittore del fatto che poi il tumore sia presente nel registro californiano

Soluzioni

a) P(A)=5235/121196=0.0432
b) P(B)=3103/121196=0.0256
c) $P(A \cap B)$=2991/121196=0.0247
d) $P(A|B)$=2991/3103=0.9639
e) $P(B|A)$=2991/5235=0.5713
f) Se=2991/3103=0.9639
g) Sp=115849/118093=0.9810

Esercizio 7

Se la media e la deviazione standard dei valori di sideremia degli uomini sani sono rispettivamente 120 e 15 microgrammi per 100 ml, qual è la probabilità che un campione casuale di 50 uomini normali abbia una media compresa tra 115 e 125 microgrammi per 100 ml?

Soluzione: $P(115 \leq \bar{x} \leq 125) = \mathrm{P}\left(\frac{115-120}{15/\sqrt{50}} \leq z \leq \frac{125-120}{15/\sqrt{50}}\right) = P(-2.36 \leq z \leq 2.36) = 0.9909 - 0.0091 = 0.9818$

Esercizio 8

Il national health and nutrition examination survey del 1988-1994 ha riscontrato un livello medio di colesterolo sierico di 180 per le donne di età 20-29 anni. La deviazione standard era approssimativamente di 37. Se estraiamo un campione di 60 donne da questa popolazione, qual è la probabilità che il livello medio del colesterolo sierico nel campione sia:

a) Fra 170 e 195
b) Sotto 175
c) Maggiore di 190

Soluzioni

a) $P(170 \leq \bar{x} \leq 195) = P\left(\frac{170-180}{37/\sqrt{60}} \leq z \leq \frac{195-180}{37/\sqrt{60}}\right)$
$= P\left(\frac{-10}{4.78} \leq z \leq \frac{15}{4.78}\right) = P(-2.09 \leq z \leq 3.14)$
$= 1 - 0.01831 - 0.00097 = 0.9807$

b) $P(\bar{x} \leq 175) = P\left(z \leq \frac{175-180}{37/\sqrt{60}}\right) = P\left(z \leq \frac{-5}{4.78}\right)$
$= P(z \leq -1.05) = 0.14686$

c) $P(\bar{x} \geq 190) = P\left(z \geq \frac{190-180}{37/\sqrt{60}}\right) = P\left(z \geq \frac{10}{4.78}\right)$
$= P(z \geq 2.09) = 0.01831$

Esercizio 9

Vogliamo stimare il numero medio di battiti cardiaci al minuto per una certa popolazione. Il numero medio di battiti al minuto per un campione di 49 soggetti era risultato pari a 90. Assumiamo che questi 49 pazienti costituiscano un campione casuale e che la popolazione sia distribuita normalmente con una deviazione standard di 10. Costruire l'intervallo di confidenza al 90%, 95% e 99%.

Soluzione

- $IC_{90\%} = \bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n} = 90 \pm 1.64 \cdot 10/\sqrt{49} = 90 \pm 1.64 \cdot 1.43 = [87.65; 92.35]$
- $IC_{95\%} = \bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n} = 90 \pm 1.96 \cdot 10/\sqrt{49} = 90 \pm 1.96 \cdot 1.43 = [87.20; 92.80]$
- $IC_{99\%} = \bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n} = 90 \pm 2.58 \cdot 10/\sqrt{49} = 90 \pm 2.58 \cdot 1.43 = [86.31; 93.69]$

Esercizio 10

Oexle et al. hanno calcolato il VPN per un test per i portatori di deficit X-linked di ornitina transcarbamilasi (OTCD, un disturbo del ciclo dell'urea). Il test, noto come «Test allopurinolo», è spesso utilizzato per individuare i potenziali portatori di OTCD. Questoi ricercatori citano Brusilow e Horwich per la stima della sensibilità del test allopurinolo, pari a 0.927, e hanno ricavato la specificità, ottenendo un valore di 0.997. A questo punto, avendo stimato una prevalenza di OTCD nella popolazione generale pari a 1/32000, hanno utilizzato il teorema di Bayes per calcolare il VPN di questo test di screening. Quale valore hanno ottenuto?

Soluzione

Se=0.927
Sp=0.997
P=1/32000=

VPN=P(S|T-)=?

$$P(S|T^-) = \frac{P(S \cap T^-)}{P(T^-)} = \frac{P(T^-|S) \cdot P(S)}{P(T^- \cap S) + P(T^- \cap M)}$$

$$= \frac{P(T^-|S) \cdot P(S)}{P(T^-|S) \cdot P(S) + P(T^-|M) \cdot P(M)} = \frac{Sp \cdot (1-p)}{Sp \cdot (1-p) + (1-Se) \cdot p}$$

$$= \frac{0.997 \cdot (1 - 1/32000)}{0.997 \cdot (1 - 1/32000) + (1 - 0.927) \cdot 1/32000}$$

$$= \frac{0.9969688438}{0.9969688438 + 0.0000022813} = 0.9999977118$$

Esercizio 11

Un dietologo di una ASL vuole condurre un'indagine fra una popolazione di ragazze adolescenti per determinare il consumo proteico medio giornaliero in grammi. Il dietologo desidera ottenere un intervallo di confidenza ampio circa 10gr, con una confidenza del 95%. Dall'esperienza passata è noto che la deviazione standard della popolazione è di circa 20gr.

Quante ragazze dovrà campionare il dietologo?

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{\Delta}\right)^2 = \left(\frac{1.96 \cdot 20}{5}\right)^2 = 7.84^2 = 61.47 \cong 62$$