FACOLTA' DI FARMACIA
Corso di Laurea in Scienza del Farmaco
I Prova mensile di Matematica e Informatica
A.A. 2011-2012

Tema A

Cognome e nome..

Esercizio n. 1.- A un gruppo di 1000 consumatori di un supermercato è stato chiesto quale tipo di jogurt avevano acquistato (A, B, C).
E risultato che 650 avevano acquistato A, 300 il B e 400 C. Inoltre 90 hanno dichiarato che le avevano acquistate tutte e tre, 240 A e B, 210 A ed C, 120 B e C.. Trovare il numero di consumatori corrispondente a ciascun insieme compreso quello che non ha acquistato alcun jogurt.
Dare la definizione di distribuzione di probabilità ed individuare la distribuzione collegata con l'esperimento precedente.

Svolgimento

Il roblema ci propone tre insiemi A, B ed C che si intersecano parzialmente e sono parte di un insieme U di 1000 soggetti sottoposti ad un test per sulla loro preferenza. Tralasciando per comodità il segno di intersezione abbiamo: n(ABC) = 90, n(AB) = 240, n(AC) = 210, n(BC) = 120, per un totale di n(A)= 650, n(B) = 300, n(C) = 400.

$(A\cup \bar{A})\cap(B\cup \bar{B})\cap (C\cup \bar{C}) = (AB\cup A\bar{B} \cup \bar{A} B\cup \bar{A}\bar{B}) \cap (C\cup \bar{C}) = ABC\cup AB\bar{C} \cup A\bar{B}C\cup A\bar{B}\bar{C} \cup \bar{A}BC\cup$

$\bar{A}B\bar{C} \cup \bar{A}\bar{B}C\cup \bar{A}\bar{B}\bar{C} = U$

Essi costituiscono una partizione dell'insieme U.
Per calcolare il numero di elementi di ciascun insieme (o evento) possiamo costruire un diagramma di Venn notando che:

$ABC=90$; $AB\bar{C}$ =240; $A\bar{B}\bar{C}$ =230; $AB\bar{C}$ =150; $\bar{A}BC$=30; $\bar{A}B\bar{C}$ =30; $\bar{A}\bar{B}C$ =100; $\bar{A}\bar{B}\bar{C}$ =190

Considerando come **evento** l'appartenenza a ciascuno di questi insiemi possiamo costruire 8 eventi possibili con le loro probabilità date dalle frequenze relative.

$P\{ABC\}$ = 90/1000 = 0,09 = p_1

$P\{AB\bar{C}\}$ = 150/1000 = 0,15 = p_2

$P\{A\bar{B}C\}$ = 180/1000 = 0,18 = p_3

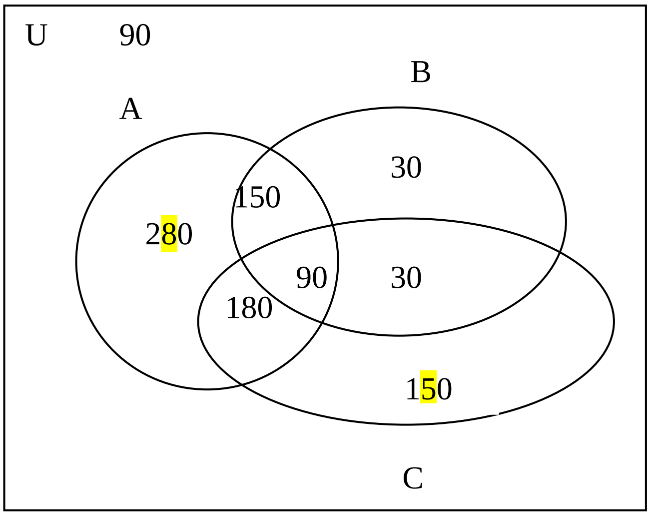
$P\{A\bar{B}\bar{C}\}$ = 280/1000 = 0,28 = p_4

$P\{\bar{A}\bar{B}C\}$ = 30/1000 = 0,03 = p_5

$P\{\bar{A}B\bar{C}\}$ = 30/1000 = 0,03 = p_6

$P\{\bar{A}\bar{B}C$ = 100/1000 = 0,15 = p_7

$P\{\bar{A}\bar{B}\bar{C}\}$ = 190/1000 = 0,09 = p_8



Gli 8 eventi sono a due a due incompatibili e la loro somma è un evento certo. Con le loro probabilità P_K formano una distribuzione di probabilità.

Infatti : 0,09 + 0,15 + 0,18 + 0,28 + 0,03 + 0,03 + 0,15 + 0,09 = 1

p_1 p_2 p_3 p_4 p_5 p_6 p_7 p_8

Esercizio n. 2.- Un gruppo di 200 pazienti della clinica psichiatrica si è lamentato di non riuscire a dormire bene. Ad alcuni di essi vengono somministrate delle pillole di sonnifero mentre ad altri vengono date delle pillole di zucchero (benché sia stato detto loro che si trattava di sonnifero). E' stato chiesto in seguito se le pillole siano state utili o meno. I risultati delle risposte sono riportate nella tabella seguente:

Sonnif.\Dorme	Bene	Male
SI	58	22
NO	50	70

Supponendo che tutti i pazienti abbiano detto la verità
a) determinare se si evidenzia una correlazione tra sonnifero e sonno secondo il teorema di Bayes, e provare l'ipotesi che non ci sia differenza tra le pillole di sonnifero e quelle di zucchero al livello di significatività di 0,05
b) indicare le distribuzioni di probabilità che intervengono nel procedimento.

Svolgimento
Completiamo la tabella data con i totali marginali:

Sonnif\Dorme	Bene	Male	Totale
Si	58	22	80
No	50	70	120
Totale	108	92	200

Dividendo ciascun valore per il totale generale otteniamo le frequenze relative che consideriamo come probabilità dei rispettivi eventi:

Sonnif\Dorme	Bene	Male	Totale
Si	0,29	0,11	0,4
No	0,25	0,35	0,6
Totale	0,54	0,46	1

Per determinare il coefficiente di correlazione tra le due caratteristiche (sonnifero e sonno) bisogna la probabilità condizionata che un paziente che ha preso il sonnifero dorma bene e confrontarla con la probabilità generale che un paziente dorma bene.
Per il teorema sulle probabilità condizionate abbiamo infatti la relazione
$P\{S \wedge B\} = P\{B\}\, P\{S/B\}$, dalla quale si ricava: $P\{S/B\} = P\{S \wedge B\}/P\{B\} = 0{,}29/0{,}54 = 0{,}537$
Dividendo questo valore per la probabilità totale di dormire bene si calcola il coefficiente di correlazione:

$$\rho = \frac{P\{S/B\}}{P\{S\}} = \frac{0{,}537}{0{,}4} = 1.34$$

Essendo questo valore maggiore di 1 si può ipotizzare che il sonnifero sia efficace.
b) Nel problema sono intervenute tre distribuzioni di probabilità. Quella principale con 4 eventi $S \wedge B$, $N \wedge B$, $S \wedge M$, $N \wedge M$ con probabilità rispettivamente di: 0,29, 0,25, 0,11 e 0,35.

Le altre due sono le **distribuzioni marginali** di tipo bernoulliano con due eventi (S, N) e probabilità P{S} = 0,4 e P{N} = 0,6. L'altra avrà anche due eventi (B, M) con probabilità P{M} = 0,46 e P{B} = 0,54.

Facoltativo. Verifichiamo l'ipotesi H_0 che le due caratteristiche siano indipendenti. In questo caso avremo che il 50% dei pazienti che ha preso il sonnifero (B) avrà dormito bene, gli altri 50% male. Lo stesso per quelli che non hanno preso il sonnifero (N) riassunto nella seguente tabella.

Sonnif\Dorme	Bene	Male	Totale
Si	54	46	100
No	54	46	100
Totale	108	92	200

Confrontando le due tabelle con il test del χ^2si ottiene:
$\chi^2 = (58 - 54)^2/54 + (22-46)^2/46 + (50-54)^2/54 + (70-46)^2/46 = 25,6$
Essendo questo valore superiore al valore critico per un grado di libertà $(n_r - 1)(n_c - 1) = 1$
Dobbiamo rifiutare l'ipotesi che le due caratteristiche siano indipendenti e assumere quella alternativa che siano correlate.

Esercizio n. 3.- Se la distribuzione C(x) del colesterolo di un individuo sano è normale con media μ = 170 ml/dl e scarto quadratico medio s = 26 ml/dl calcolare i valori standard per le quantità C_1= 115, C_2 =145, C_3 = 210 e calcolare la probabilità che sia C>210 ml/dl oppure C<115ml/dl.
Svolgimento
I valori della variabile aleatoria che segue una distribuzione normale di speranza matematica μ scarto quadratico medio s si ottengono dalla formula di Gauss mediante la formula che la definisce:

$z = (x - \mu)/s$. Pertanto $z_1 = (115 – 170)/26 = -2,115$, $z_2 = (145 – 170)/26 = -0,96$ e
$z_3 = (210 – 170)/26 = 1,528$.

La probabilità che C sia minore di 115 si legge sulla tabella della funzione di ripartizione F(z). Tenendo conto che le tabelle presentano i valori della F(z) da zero in poi e quindi l'area sotto la curva normale da 0 all'infinito è uguale a 0,5, tenendo anche conto che la funzione densità di probabilità f(z) è simmetrica rispetto all'origine, i valori trovati sulle tabelle devono venir aumentati di 0,5 per i valori di z positivi, devono venir sottratti da 0,5 per i valori di z negativi.
L'area sotto la curva normale da $-\infty$ a -3.15 sarà data da 0,5 il valore letto in tabella per z positivo.
La probabilità cercata è data da
P{C<115} = 0,5 - F(2,115). Perciò: P{C<115} = 0,5 - 0,4828 = 0,0172
Analogamente per P{C>210} bisogna sottrarre da 0,5 (area totale da 0 a $+\infty$) il valore della F(-0,96).
Quindi P{C>210}= 0,5-0,3315 =0,1685

Esercizio n. 4.- Il valor medio del colesterolo di un campione di 100 degenti della Clinica medica è di 165 ml/dl e lo scarto quadratico medio è di 24 ml/dl.
Determinare l'intervallo di confidenza al 90%, 95%e 99% per la media della popolazione.

Svolgimento

L'intervallo di confidenza per la media di un campione di 100 elementi è dato dal valore critico di z per il livello di confidenza indicato, moltiplicato per lo scarto quadratico medio del campione e diviso per la radice quadrata del numero di elementi del campione. In simboli:

$z_c \dfrac{s}{\sqrt{n}}$ dove n=100, s = 24.0. E quindi se z_{90}= 1,65; z_{95}=1,96 e z_{99} =2 ,58 troviamo:

Al livello di confidenza del 90%: 165 ± 1,65*24/10 = 165 ± 3,96
Al livello di confidenza del 95%: 165 ± 1,96*24/10 = 165 ± 4,70
Al livello di confidenza del 99%: 165 ± 2,58*24/10 = 165 ± 6,19

Per trovare la numerosità del campione che assicuri un intervallo minore di 3mg/ml poniamo:

$z_c \dfrac{s}{\sqrt{n}} = 3$ da cui ricaviamo $z_c \dfrac{s}{3} = \sqrt{n}$ dalla quale troviamo n = $\left(z_c \dfrac{s}{3} \right)^2$ e di conseguenza:

per z_c = 1,96 sarà n = 284; per z_c=1,65 sarà n = 201 e per z_c=2,58 si ottiene n = 492.

Esercizio n. 5.- Se la probabilità che un individuo sia affetto da talassemia è 0,002. Determinare la probabilità che su 10000 individui: a) esattamente 1, b) più di 2 siano talassemici.
(è sufficiente indicare le formule senza eseguire i calcoli).

Svolgimento

Poiché la probabilità che la patologia si presenti è molto piccola (0,002) possiamo dire di trovarci di fronte ad **eventi rari**. Poiché, a quanto se ne sa gli eventi sono anche indipendenti possiamo pensare di applicare la distribuzione di Poisson con una media ipotetica che possiamo derivare dalla distribuzione binomiale: m = n*p = 10000*0,002 = 20.

Nella distribuzione di Poisson la probabilità di ottenere k successi è data da: $p_k = \dfrac{m^k e^{-m}}{k!}$.

Di conseguenza possiamo calcolare la probabilità di ottenere 0, 1, 2 successi:

$p_0 = \dfrac{m^0 e^{-m}}{0!}$; $p_1 = \dfrac{m^1 e^{-m}}{1!}$; $p_2 = \dfrac{m^2 e^{-m}}{2!}$

Pertanto la probabilità di trovare più di due individui affetti da questa patologia sarà:

$$P\{k>2\} = 1 - (p_0 + p_1 + p_2).$$