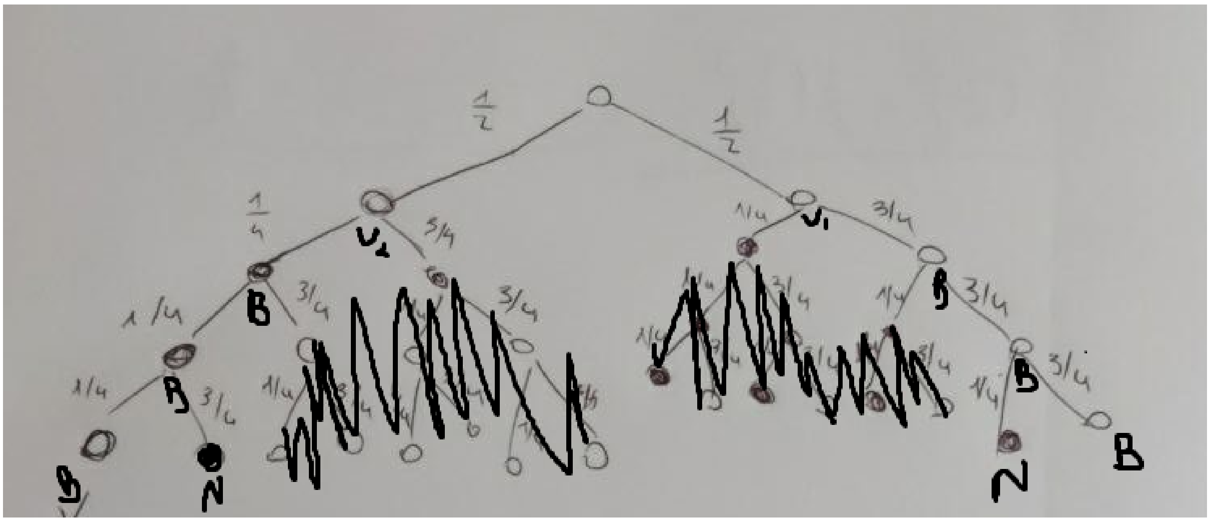
Esame CPS Aprile 19

1) Nell'urna 1 le palline bianche sono il triplo delle nere, nell'urna 2 le palline nere sono il triplo delle bianche. Scegliamo a caso un'urna e eseguiamo estrazioni con reimmissione della pallina estratta dall'urna scelta. Sapendo che le prime due estrazioni hanno dato entrambe pallina bianca, con che probabilità l'estrazione successiva darà ancora pallina bianca?
a. 5/8 b. 7/10 c. ¾ d. ⅗

Svolgo il problema con casi favorevoli diviso casi totali, rappresento tramite alberi gli esiti possibili e metto al numeratore tutti gli esiti che verificano il problema e al denominatore tutti quelli esistenti, avendo però la condizione che sono uscite già due palline bianche.



Se ora prendo tutti i casi favorevoli ovvero la probabilità che capiti 3 bianche e sommo gli esiti delle due urne, e divido per i casi totali, ovvero tutti i casi che la terza pallina sia bianca sommati a quelli che sia nera per entrambe le urne e ottengo

$$\frac{\frac{1}{2}(\frac{1}{4})^3 + \frac{1}{2}(3/4)^3}{\frac{1}{2}(\frac{1}{4})^3 + \frac{1}{2}(3/4)^3 + \frac{1}{2}(3/4)(\frac{1}{4})^2 + \frac{1}{2}(\frac{1}{4})(3/4)^2} = 7/10$$

2) Uno studente viene sottoposto a 10 domande a risposta multipla con 4 opzioni di risposta, scelte a caso indipendentemente da un certo database. Tra queste, 8 sono di probabilità e 2 di statistica. Lo studente conosce la risposta corretta a 5/8 delle domande di probabilità, ma nessuna alle domande di statistica. Di fronte ad una domanda di cui non conosce la risposta risponde a caso. Qual è il numero di atteso di domande risposte correttamente?
a. 25/4 b. 5 c. 15/2 d. 11/2

Consideriamo che delle 8 domande di probabilità lo studente ne sa rispondere a sicuramente a 5, quindi ne restano fuori 3 a cui non sa rispondere e queste le butta a caso con P=1/4, stessa cosa per quelle 2 di statistica, quindi il valore atteso è :

E(x) = 5*1 + (8-5)* ¼ + 2 * ¼ = 5 + 5/4 = 25/4

3) Se X e Y sono variabili aleatorie esponenziali indipendenti di media 1, quale tra le seguenti funzioni, definite per x>0, è proporzionale alla densità di X+Y?
a. exp{-x/2} b. x^2exp{-x} c. exp{-x^2/8} d. xexp{-x}

La somma di due variabili aleatorie esponenziali è una variabile aleatoria Gamma di parametri (k,λ2) con k uguale alla somma delle medie e λ2 = λ1
E[X] = 1/λ1 = 1 => λ1 = 1 . E[Y] = 1/λ1 = 1 => λ1 = 1. Quindi k=2 e λ2 = 1.
Una variabile Gamma è direttamente proporzionale a
x^(k-1)e^(- λx) => x^(2-1) e^(-x) = xe^(-x)

4) Sia S la somma di 12 variabili aleatorie indipendenti, uniformi nell'intervallo (0,1). Utilizzando una celebre disuguaglianza la probabilità che S prenda valori tra 4 e 8 è almeno
a. 9/16 b. 15/16 c. ¾ d. ½

Dato che le variabili aleatorie sono indipendenti, sappiamo che nell'intervallo (a,b)
E[X1] = E[X2] = ... = E[X12] = (b-a)/2 = (1-0) / 2 = ½
Var(X1) = Var(X2) = … = Var(X12) = (b-a)^2 /12 = (1-0)^2 /12 = 1/12
E[S] = sommatoria (i=1 a 12) E[Xi] = 12 * 1/2 = 6
Var(S) = sommatoria (i=1 a 12) Var[Xi] = 12 * 1/12 = 1
Per Chebyshev
P(|S-μ|/σ <= c) >= 1 - 1/c^2 => P (μ-σx <= S <= μ+σc) >= 1 - 1/c^2 =>μ-σc = 4 => 6 - c = 4 => c = 2
Quindi 1 - 1/c^2 = 3/4

5) Si consideri una densità nell'intervallo (-π,π) proporzionale a |cos x|. Qual'è la differenza interquartile?
a. 2π b. √2 c. π/2 d. π

$$f(x) = k\,|\cos(x)| \Rightarrow \int_{-\pi}^{\pi} k\,|\cos(x)|\,dx = 1 \Rightarrow$$
$$\Rightarrow \int_{-\pi}^{0} k\cos(x)\,dx + \int_{0}^{\pi} k\cos(x)\,dx = 1 \Rightarrow 4k = 1 \Rightarrow k = \frac{1}{4}$$

La differenza interquartile è la differenza tra il terzo quartile e il primo quartile e quindi:
Primo Quartile

$$\int_{-\pi}^{a} \frac{1}{4}\cos(x)\,dx = \frac{1}{4} \Rightarrow \frac{\mathrm{sen}(a)}{4} = \frac{1}{4} \Rightarrow \cancel{a} - \mathrm{sen}(a) = 1 \Rightarrow a = -\frac{\pi}{2}$$

Terzo quartile

$$\int_{b}^{\pi} \frac{1}{4}\cos(x)\,dx = \frac{1}{4} \Rightarrow \frac{\mathrm{sen}(b)}{4} = \frac{1}{4} \Rightarrow \mathrm{sen}(a) = 1 \Rightarrow a = \frac{\pi}{2}$$

e quindi $\frac{\pi}{2} + \frac{\pi}{2} = \pi$

6) Se X ha una densità normale con varianza 2 e media ignota, determinare quante osservazioni n indipendenti occorre fare su X in modo che l'intervallo di confidenza, centrato sulla media campionaria di livello di confidenza 0.95, abbia lunghezza ½ . Evitando di consultare le tavole, si assuma che, se Z è una variabile aleatoria normale standard, $P(|Z|<2)=0.95$.
a. 16 b. 128 c. 64 d. 32

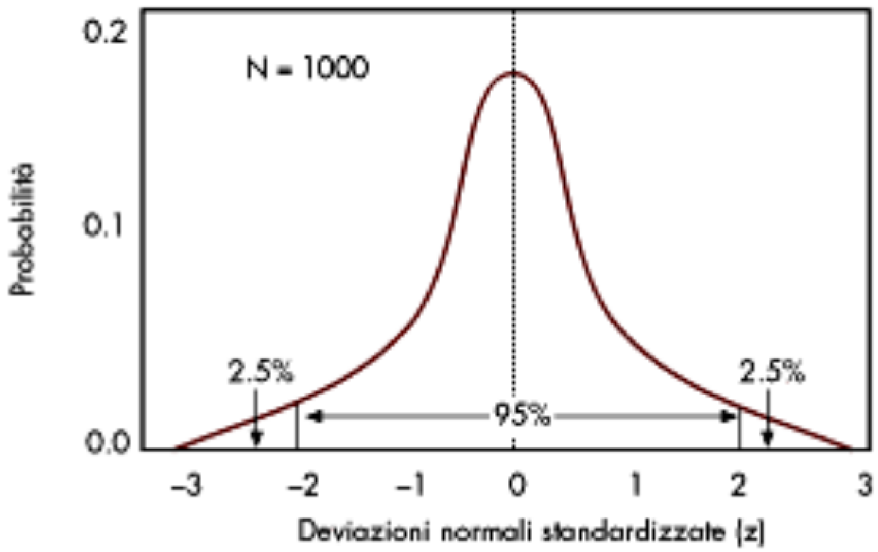
Se scelgo un n grande a piacere (per il teorema del limite centrale), posso approssimare X ad una normale di parametri mu,2

$X \sim N(\mu, 2)$, L'intervallo di confidenza é: $\mu \in \left(\bar{X} - \frac{6}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}}, \bar{X} + \frac{6}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}}\right) \rightarrow$

$\rightarrow$ La distanza é: $\cancel{\bar{X}} + \frac{6}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}} - \cancel{\bar{X}} + \frac{6}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}} = \frac{1}{2} \rightarrow$

$\rightarrow 2 \cdot \frac{6}{\sqrt{n}} \cdot z_{\frac{\alpha}{2}} = \frac{1}{2} \rightarrow z_{\frac{\alpha}{2}} = 2 \rightarrow \sqrt{n} = 46 \cdot z_{\frac{\alpha}{2}} \rightarrow$

$\rightarrow \sqrt{n} = 4 \cdot 2\sqrt{2} \rightarrow n = 16 \cdot 4 \cdot 2 = 128$



7) Il gioco del memory consta di 8 tessere, divise in 4 gruppi da 2, con etichetta differente per ciascun gruppo, es. 2 fragole, 2 pere, 2 banane e 2 mele, che vanno sistemate in 8 distinte posizioni. Se si considerano equivalenti due configurazioni in cui si permutano i 4 tipi di etichetta, quante sono le possibili configurazioni?
a. 2520 b. 70 c. 105 d. 1680

In questo esercizio abbiamo delle disposizioni semplici di k=4 coppie di elementi distinti. Per ogni k elemento sappiamo anche che non ci interessa la permutazione della stessa coppia, quindi è una disposizione semplice che si risolve con

$$D_{n,k} = n! / (n-k)! = 8! / 4! = 1680$$

8) Se X, Y e Z sono variabili aleatorie standard con tutti i coefficienti di correlazione tra coppie di variabili uguali, indicare rispettivamente il valore minimo e il valore massimo della varianza di X+Y-Z.

a. 0 e 4 b. 1 e 5 c. 1 e 3 d. 3 e 5

Var(X+Y-Z) = Var(X) + Var(Y) + Var(Z) + 2Cov(X,Y) - 2 Cov(X,Z) - 2 Cov(Y,Z)
Var(X) = Var(Y) = Var(Z) = 1

$$Cov(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)\cdot Var(y)}}$$

so che il denominatore è uguale a
Corr(X,Y) = Cov(X,Y) = Cov(X,Z) = Cov(Y,Z)
Var(X+Y-Z) = 3 - 2 Cor(Y,Z)
-1 <= Corr(X,Y) = Corr(X,Z) = Corr(Y,Z) <= 1
Var(X+Y-Z)max = 3 - 2 (1) = 1,
Var(X+Y-Z)min = 3 - 2 (-1) = 5.

9) Se i segni 1, X e 2 di ciascuna partita di calcio sono equiprobabili, con che probabilità in 4 partite indipendenti l'una dall'altra sono presenti tutti e 3 i segni?

a. tutte le altre risposte proposte sono errate b. ⅓ c. ¼ d. 4/9

1 * ⅓ * ⅓ * 1 (4 3) = 4/9
1 = prob che esca uno dei 3 esiti (1,X,2) = 1
⅓ = prob che esca uno tra (S-primo esito)= 2
⅓ = prob che esca (S- primo esito - secondo esito) = X
1 = prob che alla quarta pescata esca qualsiasi cosa = (1,X,2)

10) Sia X una variabile aleatoria positiva di media finita e F un evento indipendente da X, con P(F)= 10^{-3}. Definita Y=10^{5}X1F, dove 1F è l'indicatrice dell'evento F, dire quanto valgono, rispettivamente, P(Y>X) e E(Y)/E(X).

a. 10-5, 102 b. 10-2 103 c. 10-3, 105 d. 10-3, 102

1F é la funzione indicatrice : $\begin{cases} 1F = 1 & \text{se } x \in F \\ 1F = \emptyset & \text{se } x \notin F \end{cases}$

$P(x \in F) = 10^{-3} \longrightarrow Y = \begin{cases} 10^5 X & \text{se é vero F} \\ \emptyset & \text{se non é vero F} \end{cases}$

P(Y>X) Y é maggiore di X se e solo se $Y = 10^5 X \rightarrow$ solo se é meno F $\rightarrow$

$\rightarrow P(Y>X) = 10^{-3}$

$E[Y] = 10^2 E[X] \cdot 10^{-3} \longrightarrow \frac{E[Y]}{E[X]} = \frac{10^5 \cdot 10^{-3} \cancel{E[X]}}{\cancel{E[X]}} = 10^2$