Probability and Mathematical Statistics: Homework #4

Due on October 22, 2022 at 11:59am

0

Name: **Zhu Zhelin**
Student ID: 2021533077

Problem 1 (mention the source of question, *e.g.*, BH CH0 #1)

Solution

First I want to calculate E(X)

$$E(X) = \sum_{k=1}^{\infty} \frac{cp^k}{k} \cdot k = \sum_{k=1}^{\infty} cp^k = \frac{cp}{1-p}$$

to calculate $Var(X)$,we can calculate $E(X^2) - (EX)^2$

$$E(X^2) = \sum_{k=1}^{\infty} \frac{cp^k}{k} \cdot k^2 = c\sum_{k=1}^{\infty} kp^k = cp\sum_{k=1}^{\infty} \frac{\mathrm{d}p^k}{\mathrm{d}p} = cp\frac{\mathrm{d}\sum_{k=1}^{\infty} p^k}{\mathrm{d}p} = cp \cdot \frac{\mathrm{d}\frac{p}{1-p}}{\mathrm{d}p} = \frac{cp}{1-p^2}$$

, so the answer is $\frac{cp(1-cp)}{(1-p)^2}$

Problem 2 (BH CH0 #2)

Solution

(a) We can let I_i be the indicator random variable represent that in the i_{th} trial,show if two people simultanously success,$P(I_i) = p_1p_2, P(I_i^c) = 1 - p_1p_2$, and I_i is $i.i.d$ satisfy $I_i \sim Bern(p_1p_2)$,then I satisfy $Bin(n, p_1p_2)$,then I can define $Z \sim Geom(p_1p_2)$ represent the distribution of the Bern trials,Z is the number of faiure trial number before first success happen,then $Y = Z + 1$ satisfy the distribution of first success happen,so the distribution is $P(Y = k) = P(Z + 1 = k) = P(Z = k - 1) = (1 - p_1p_2)^{k-1}p_1p_2$ the expected time is $E(Y) = E(Z + 1) = E(Z) + 1 = \frac{1}{p_1p_2}$

(b) the same as below,we can define I_i as at least one success happend at i_{th} trial ,$P(I_i^c) = (1 - p_1)(1 - p_2), P(I_i) = 1-(1-p_1)(1-p_2)$,$I_i$satify $\sim$Bern$(1-(1-p_1)(1-p_2))$4then the number of trials(including first success) until at least one success happen represented by Y,then $Z = Y - 1 \sim$ Geom$(1-(1-p_1)(1-p_2))$ $E(Y) = E(Z + 1) = \frac{1}{1-(1-p_1)(1-p_2)}$

(c) considering the random variable X be the time in k_{th} trial,they simultaneously and firstly win,when it happens,it should satisfy that $(k - 1)$trials they both failure the possibilty of both failure is $(1 - p_1)(1-p_2) = (1-p_1)^2$,then P(X=K)=$(1-p_1)^{2(k-1)}p_1^2$,then sum all the circumstances up ,I can get that $\sum_{k=1}^{\infty}(1-p_1)^{2(k-1)}p_1^2 = \frac{p_1}{2-p_1}$,then according to symmetry,the probability of Nick's first success precedes Penny is equal to Penny's first success preccedes Nick's success ,we can get the probability by calculate $\frac{1}{2}(1 - \frac{p_1}{2-p_1}) = \frac{1-p_1}{2-p_1}$

Problem 3 (BH CH0 #3)

(a) We can first define X as the different number of stops it will happen among k people and X= $I_2+I_3+\cdots I_n$ among them

$I_j = 1$(when j_{th} floor stop),I_j =0,other,$E(I_j) = 1 - P$(nobody stop at j_{th}floor)=1-$\left(\frac{n-2}{n-1}\right)^k$ $E(X) = E(I_2 + I_3 + \cdots I_n) = (n-1)\cdot(1-(1-\frac{1}{n-1})^k)$

(b) the same as below ,but the $E(I_j)$ changes to be $1-(1-p_j)^k$,so the total probability is $\sum_{j=2}^{n} 1-(1-p_j)^k = n-1-\sum_{j=2}^{n}(1-p_j)^k$

Problem 4 (BH CH0 #4)

Solution

(a) the LOTUS define that $E(g(x)) = \sum_x g(x)P(X = x)$,to calculate the lefthandside I can let $h(X) = Xg(X)$,then

$$E(h(X)) = \sum_k kg(k)P(X = k) = \sum_{k=0} kg(k) \cdot \frac{e^{-\lambda}\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} g(k) \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} = \lambda \sum_{k=0}^{\infty} g(k+1) \frac{e^{-\lambda}\lambda^k}{k!}$$

and the rhs

$$\lambda E(g(X+1)) = \lambda \sum_{k=0}^{\infty} g(k+1) \frac{e^{-\lambda}\lambda^k}{k!} = lhs$$

,so Q.E.D

(b)

$$E(X^3) = E(X \cdot X^2) = \lambda E((X+1)^2) = \lambda E(X^2+2X+1) = \lambda(E(X^2)+2E(X)+E(1)) = \lambda(E(X \cdot X)+2E(X \cdot 1)+1) =$$

$$\lambda(\lambda E(X+1) + 2\lambda E(1) + 1) = \lambda^3 + 3\lambda^2 + \lambda$$

Problem 5 (BH CH0 #5)

(a) First ,I will use the property of CDF that is if$m_1 \leq m_2$,then $P(Y \leq m_1) \leq P(Y \leq m_2)$ since$P(Y \leq 23) = 0.507 \geq 0.5, P(Y \geq 23) = 1 - P(Y \leq 22)$,$P(Y \leq 22) < 0.5$,so $P(Y \geq 23) \geq 0.5$,so 23 is the median,then to prove the uniquness,if $m > 23, P(Y \geq m) = 1 - P(Y \leq m-1) \leq 0.493$ doesn't satisfy $P(Y \geq m) \geq 0.5$,then if $m < 23$,P($Y \leq m) \leq P(Y \leq 22) < 0.5$,which also doesn't satisfy,so that 23 is the only median.

(b) To show the fact,we can first consider when X=j,since $X \geq j$then $I_1 \cdots I_j = 1$,$I_{j+1} \cdots I_{366} = 0$,$X = I_1 + I_2 + \cdots I_{366}$,then I want to deduce from right to left.I can define the max index j which satisfy $I_j = 1$,it indicates $X \leq j$,then all the variable $I_k, k \leq j$ must be 1 since $j \geq k$,since j is the max index ,for any $i > j$,$I_i = 0$,then we can get that $X = j$ exactly ,so max index j $I_j = 1 \Rightarrow I_1 + I_2 + \cdots I_{366} = j$ and $X = j$,it is a bijection,thus prove the equality

$$E(X) = E(I_1) + E(I_2) + \cdots E(I_{366}) = \sum_{j=1}^{366} p_j$$

(c) use python to solve it numerically,we can get the expected time equal to 24.616

(d) We want to calculate $Var(X)$,$Var(X) = E(X^2) - (EX)^2$,$E(X^2)$,first,I want to handle $X^2 = I_1^2 + I_2^2 \cdots I_{366}^2 + 2\sum_{j=2}^{366}\sum_{i=1}^{j-1} I_i I_j$,then $I_k^2 = I_k$(proved by enumerate two circumstances),$I_i I_j$when(i< j),this equalt to $I_j^2 = I_j$,so it can be reduced to$I_1 + \cdots I_{366} + 2\sum_{j=2}^{366}(j-1)I_j$,so E($X^2$)=$E(X) + 2\sum_{j=2}^{366}(j-1)p_j$,Var(X)=$\sum_{j=1}^{366} p_j - (\sum_{j=1}^{366} p_j)^2 + 2\sum_{j=2}^{366}(j-1)p_j$ using python I can calculate it numerically as 148.64

```
sum=2
p=1
for  i in range(3,367):
 p=p*(1-(i-2)/365)
 sum+=p
s=0
p=1
for i in range(2,367):
    p=p*(1-(i-2)/365)
    s+=2*(i-1)*p
v=sum-sum*sum+s
print(v)
```