Probability and Mathematical Statistics: Homework #5

Due on September 18, 2022 at 11:59am

0

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

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

Solution

$P_k = P(N = K); P_0 = 0, P_1 = 0, P_2 = 0, P_3 = p^3$,then to use first step method,I can define S_i :the result of i_{th} trial ,$S_i = 1$(H),$S_i = 0$(T),P($S_i = 1$)=p=$\frac{1}{2}$,P($S_i = 0$)=q=$\frac{1}{2}$

$$P(N = k) = P(N = k, |S_1 = 0)P(S_1 = 0) + P(N = K|S_1 = 1, S_2 = 0)P(S_1 = 1, S_2 = 0) + P(N = K|S_1 = 1, S_2 = 1, S_3 = 0) \cdot P(S_1 = 1, S_2 = 1, S_3 = 0) = P(N = K - 1)q + P(N = k - 2)pq + P(N = k - 3)p^2q \tag{1}$$

then to use PGF $g(t) = E(t^N) = \sum\limits_{k=0}^{\infty} P_k t^k = \sum\limits_{k=3}^{\infty} P_k t^k = p^2qt^3 + \sum\limits_{k=4}^{\infty} P_k t^k$
on the other hand
$\sum\limits_{k=4}^{\infty} P_k t^k = q \sum\limits_{k=4}^{\infty} P_{k-1} t^k + pq \sum\limits_{k=4}^{\infty} P_{k-2} t^k + p^2q \sum\limits_{k=4}^{\infty} P_{k-3} t^k$
$g(t) - p^2qt^3 = qt \cdot g(t) + pqt^2 \cdot g(t) + p^2qt^3 g(t)$ then I can get g(t)=$\frac{-p^2qt^3}{qt+pqt^2+p^2qt^3-1}$
E(N)=$g'(t)|_{t=1}$,since p=q=$\frac{1}{2}$,so I can get that $g'(t) = \frac{24t^2-8t^3-2t^4}{(4t+2t^2+t^3-8)^2}$,$g'(1) = 14$
Var(N)=$E(N^2) - (E(N))^2 = E(N(N-1)) + E(N) - (E(N))^2 = g''(1) + E(N) - (E(N))^2 = 142$

Problem 2 (BH CH0 #2)

(a) E(T)=E(e^{-3X})=$\sum_{k=0}^{\infty}\frac{e^{-3k}\cdot e^{-\lambda}\lambda^k}{k!}$ $=\sum_{k=0}^{\infty}\frac{e^{-\lambda}(\frac{\lambda}{e^3})^k}{k!}$ $=e^{(\frac{1}{e^3}-1)\lambda}$,since it is not equal to $e^{-3\lambda}$,so it isn't unbiased.

(b) To show is to calculate $E(g(X))=\sum_{k=0}^{\infty}\frac{(-2)^k e^{-\lambda}\lambda^k}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(-2\lambda)^k}{k!}=e^{-3\lambda}=\theta$
,so $E(g(X))-\theta=0$,
so it turns out to be unbiased

(c) The estimator g(X) could be negative and will be very large in terms of absolute value if X is large,,while θ is a positive small number ,so we can change it to a function $h(X)$

$$h(X)=\begin{cases}1 & X=2k(k\in N)\\ 0 & X=2k+1(k\in N)\end{cases}$$

since $\theta\leq 1$,it is strictly greater.

Problem 3 (BH CH0 #3)

(a) To find the expected number ,we first define X as the number of times all sixes have achieved all sixes and we can use indicator r.v. I_j to represent if the j_{th} roll satisfy all sixes ,the trial satify a possibility $p = (\frac{1}{6})^n$,it satisfy X$\sim$ $Binom(4 \cdot 6^{n-1}, (\frac{1}{6})^n)$,so E$(X) = 4 \cdot 6^{n-1} \times (\frac{1}{6})^n = \frac{2}{3}$

(b) when n is a very large number ,the $(\frac{1}{6})^n$ will turn to be very small,but among the $4 \cdot 6^{n-1}$ trials ,the toatl probability $\frac{2}{3}$,$X \sim Binom(4 \cdot 6^{n-1}, (\frac{1}{6})^n)$,so i can use X $Pois(\frac{2}{3})$,P(X$\geq$ 1)=1-P(X=0)=1-$e^{-\lambda} = 1 - e^{-\frac{2}{3}}$

(c) It will not change ,since the linearity of expectation doesn't change,we can easily conclude that $E(X) = E(I_1 + I_2 + \cdots)$can be divided into $E(I_1) + E(I_2) + \cdots E(I_n)$ though the variable is dependent ,the expectation doesn't change,so the answer is same.

Problem 4 (BH CH0 #4)

(a) considering when X=k,it is same as randomly choose m+k without replacement and the last time we get the last element as the labeled,the last time the probability we get the labeled is $\frac{n-m+1}{N-m-k+1}$,the probability to get (m-1) labeled in choosing m+k elements,is $\frac{\binom{n}{m-1}\cdot\binom{N-n}{k}}{\binom{N}{m+k-1}}$ so the total probability is $P(X=k)=\frac{\binom{n}{m-1}\cdot\binom{N-n}{k}}{\binom{N}{m+k-1}}\cdot\frac{n-m+1}{N-m-k+1}$
we can easily get that Y=X+m so $P(Y=k)=P(X=k-m)=\frac{\binom{n}{m-1}\cdot\binom{N-n}{k-m}}{\binom{N}{k-1}}\cdot\frac{n-m+1}{N-k+1}$

(b) there are n labeled elks,so we can use the labeled elks as indicator variables in the N trials,just as I_j : whether this capture happens between the elks labeled j and j-1,and the probability I can get is $\frac{1}{n+1}$,then I can define that I_{ij} :which is the indicator variable about whether the j_{th} unlabeled elk captured between i and $i-1$ labeled elk,then the probability of $I_{ij}=\frac{1}{n+1}$,then $E(X)=E(X_1+X_2+\cdots X_n)=E(X_1)+E(X_2)\cdots=E(I_{11}+I_{12}+I_{13}\cdots)+E(I_{21}+I_{22}+\cdots)\cdots=\sum_{i=1}^{m}\sum_{j=1}^{N-n}\frac{1}{n+1}=\frac{m(N-n)}{n+1}$
$E(Y)=E(X+m)=\frac{mN+m}{n+1}$

(c) since the sample is fixed,I can represent E(Y)=s,supposing Z as the number of labeled elks in the sampling Z$\sim Hgeom(n,N-n,s)$,then first I want to calculate E(Z)
$E(Z)=\sum\frac{k\binom{n}{k}\binom{N-n}{s-k}}{\binom{N}{s}}=\sum\frac{n\binom{n-1}{k-1}\binom{N-n}{s-k}}{\binom{N}{s}}=n\frac{\binom{N-1}{s-1}}{\binom{N}{s}}=\frac{nE(Y)}{N}=\frac{(nN+n)m}{nN+N}$ it is less than m,(specially when N is very large,these two are very close)

Problem 5 (BH CH0 #6)

(a) I can seperate it into two parts during the exploration phase ,you randomly eat k different dishes,so on average ,we try the dishes rank= $\frac{n+1}{2}$ the total number of ranks in this phase is $\frac{n+1}{2}k$ during the exploitation phase,you will always choose X so using LOTP ,we can get $\sum P(X=k)X(m-k)=E(X)(m-k)$,so the total is $\frac{n+1}{2}k+E(X)(m-k)$

(b) the corresponding sequence when X=j should be has the largest rank j,and the other $k-1$ dishes rank smaller than j,which means pick from $1-(j-1)$so the PMF P(X=k)=$\frac{\binom{j-1}{k-1}}{\binom{n}{k}}$

(c) first,I want to use two equation

   i j$\binom{j-1}{k-1}$=$k\binom{j}{k}$ using choosing captain or choosing monitor story telling to prove

   ii $\sum\limits_{j=k}^{j=n}\binom{j}{k}=\binom{n+1}{k+1}$ considering a queue with n+1 people with diffrent ages,choosing k+1 from n+1 is equal to considering choosing j+1 oldest person first,than choose the k people smaller than him.

   using following equation,E(X)=$\sum\limits_{j=k}^{j=n}\frac{\binom{j-1}{k-1}j}{\binom{n}{k}}=\sum\limits_{j=k}^{j=n}\frac{\binom{j}{k}k}{\binom{n}{k}}=\frac{k(n+1)}{k+1}$

(d) we can represent the whole function as $f(k)=\frac{n+1}{2}k+\frac{k(n+1)(m-k)}{k+1}$ then to solve the problem $f'(k)=0$ I can get that$(k+1)^2=2(m+1)$,since k is larger than zero,so $k=\sqrt{2m+2}-1$