Probability & Statistics for EECS: Homework #04

Due on Nov 5, 2023 at 23:59

Name: **Runkang Yang**
Student ID: 2022533080

Problem 1

1.

$$
\begin{aligned}
E(X) &= \sum_{x=1} xP(X) \\
&= 1 \times \frac{cp^1}{1} + 2 \times \frac{cp^2}{2} + 3 \times \frac{cp^3}{3} + ... \\
&= cp + cp^2 + cp^3 + ... \\
&= c(p + p^2 + p^3 + ...) \\
&= c \times \frac{p(1-p^n)}{1-p} \\
&= \frac{cp}{1-p} = -\frac{p}{(1-p)log(1-p)}
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \sum_{x=1} x^2P(X) \\
&= 1 \times \frac{cp^1}{1} + 2^2 \times \frac{cp^2}{2} + 3^2 \times \frac{cp^3}{3} + ... \\
&= cp + cp^2 + cp^3 + ... \\
&= c(p + 2p^2 + 3p^3 + ...) \\
&= \frac{cp}{(1-p)^2} = -\frac{p}{(1-p)^2log(1-p)}
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= E(X^2) - (E(X))^2 \\
&=
\end{aligned}
$$

Thus, we have the mean of X is

$$-\frac{p}{(1-p)log(1-p)}$$

the variance of X is

$$-\frac{p}{(1-p)^2log(1-p)} - (-\frac{p}{(1-p)log(1-p)})^2$$

Problem 2

1. (a)

The probability that Nick and Penny simultaneously successful is p_1p_2, so we can find that the event "they simultaneously successful at the same time after k times" has the First Success distribution with parameter p_1p_2; we denote this by $k \sim FS(p_1p_2)$. So we can easily find that the PMF is

$$P(k = n) = (1 - p_1p_2)^{n-1}p_1p_2$$

and the expected value is

$$E(k) = \frac{1}{p_1p_2}$$

(b)

For any independent trials, we can use 1 minus the probability of both fail to get the probability of at least one has a success, and the probability of both fail is $(1 - p_1)(1 - p_2)$, so we can find that the event "at least one has a success after m times" has the First Success distribution with parameter $1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1p_2$, denote it as $m \sim FS(p_1 + p_2 - p_1p_2)$. Then we can find the expected value is

$$E(m) = \frac{1}{p_1 + p_2 - p_1p_2}$$

(c)

This problem is different from (a) since it stresses the simultaneous success for the first time while (a) just asked about the success are simultaneous first (which implies that they may have fetch success before).

So before they success simultaneously at nth time, they both failed for $(n - 1)$ times, at each failed time, the probability of both failure is $(1 - p)^2$, and for the last time the probability that both success is p^2. Denote N_1 to be the times that Nick fetch his first success and N_2 to be the times that Pennny fetch his first success.

So we can get the probability that their first successes are simultaneous is

$$\begin{aligned}
P(N_1 = N_2) &= \sum_{n=1} P(N_1 = n)P(N_2 = n) \\
&= \sum_{n=1} ((1 - p)^2)^{n-1}p^2 \\
&= p^2 \sum_{n=1} ((1 - p)^2)^{n-1} \\
&= p^2 \times \frac{1 \times (1 - (1 - p)^2)^{n-1}}{1 - (1 - p)^2} \\
&= \frac{p^2}{2p - p^2} \\
&= \frac{p}{2 - p}
\end{aligned}$$

And we can find the probability that Nick's first success precedes Penny's is

$$\begin{aligned}
P(N_1 < N_2) = P(N_1 > N_2) \\
&= \frac{1 - P(N_1 = N_2)}{2} \\
&= \frac{1 - \frac{p}{2-p}}{2} = \frac{1 - p}{2 - p}
\end{aligned}$$

Problem 3

1. (a)

Let's assume that I_i denotes that the elevator will stop at the ith floor, which means that $I_i = 1$ if the elevator makes stop at the ith while $I_i = 0$ if not stop, then the probability that the elevator will stop at a certain floor is $P(X_i) = E(I_i)$ and our objective is to calculate $\sum_{i=2}^{n} E(I_i)$.
Since the event that each floor is choosen is independent with each other, which means the probability that elevator makes stop at is same,i.e.$P(X_1) = P(X_2) = ... = P(X_n)$, so we get

$$\sum_{i=2}^{n} E(I_i) = (n-1)E(I_2)$$

To calculate $E(I_2)$, which also means that we need to calculate $P(X_2)$, we find the event is to calculate the probability that there exist some people who presses the 2nd floor button. And we can do this by searching the probability that there are no person who presses the 2nd floor button then use 1 minus the later probability.
If there are no person who presses the 2nd floor button, it means that for the k people, they choose from 3 to n floors randomly, each person has (n-3+1)=(n-2) choices, while the total event is that they have (n-2+1)=(n-1) choices, so the probability that nobody presses the 2nd floor button is $(\frac{n-2}{n-1})^k$, then we get $E(I_2) = 1 - (\frac{n-2}{n-1})^k$
And

$$\sum_{i=2}^{n} E(I_i) = (n-1)E(I_2) = (n-1)(1 - (\frac{n-2}{n-1})^k)$$

(b) Similarly, we get

$$\begin{aligned} E(I_i) &= P(X_i = 1) \\ &= 1 - P(X_i = 0) \\ &= 1 - (1 - p_i)^k \end{aligned}$$

$$\begin{aligned} \sum_{i=2}^{n} E(I_i) &= \sum_{i=2}^{n} (1 - (1 - p_i)^k) \\ &= (n-1) - \sum_{i=2}^{n} (1 - p_i)^k \end{aligned}$$

Problem 4

1. (a) By LOTUS, we have

$$E(g(X)) = \sum_x g(X)P(X = x)$$

Thus, we have

$$\begin{aligned}
E(Xg(X)) &= \sum_{k=0}^{\infty} Xg(X)P(X = k) \\
&= \sum_{k=0}^{\infty} kg(k)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \lambda \sum_{k=1}^{\infty} g(k)\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} \\
&= \lambda \sum_{k=0}^{\infty} g(k+1)\frac{e^{-\lambda}\lambda^k}{k!} \\
&= \lambda \sum_{k=0}^{\infty} g(k+1)P(X = k) \\
&= E(g(X+1))
\end{aligned}$$

(b)

Suppose that $g(x) = x^2$

$$\begin{aligned}
E(X^3) &= E(Xg(X)) \\
&= \lambda E(g(X+1)) \\
&= \lambda E((X+1)^2) \\
&= \lambda E(X^2 + 2X + 1) \\
&= \lambda E(X^2) + 2\lambda E(X) + \lambda
\end{aligned}$$

From poisson distribution, we have

$$E(X) = \lambda, Var(X) = E(X^2) - (E(X))^2 = \lambda$$

so we get

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

Thus, we get

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

Problem 5

1. (a)

$P(X \leq m)$ means that the probability that there exist birthday match in the m people, since the probability that there are no birthday match is $\frac{365\times 364\times ...\times(365-m+1)}{365^m}$, so we can get the below probability

$$P(X \leq 22) \approx 0.48 \quad P(X \leq 23) \approx 0.51 \quad P(X \leq 24) \approx 0.54$$
$$P(X \geq 22) \approx 0.56 \quad P(X \geq 23) \approx 0.52 \quad P(X \geq 24) \approx 0.49$$

So by observation, it is obviously that 23 is the unique median of X.

(b)

From (a), we get

$$P(X \leq j) = 1 - \frac{365 \times 364 \times ... \times (365 - j + 1)}{365^j}$$

so

$$P(X \geq j) = 1 - P(X \leq j - 1) = 1 - (1 - \frac{365 \times 364 \times ... \times (365 - (j-1) + 1)}{365^{j-1}}) = \frac{A_{365}^{j-1}}{365^{j-1}}$$

Since

$$\begin{aligned} p_j &= (1 - \frac{1}{365})(1 - \frac{2}{365})...(1 - \frac{j-2}{365}) \\ &= \frac{364}{365} \times \frac{363}{365} \times ... \times \frac{365 - (j-2)}{365} \\ &= \frac{365 \times 364 \times ... \times (365 - (j-1) + 1)}{365^{j-1}}) \\ &= \frac{A_{365}^{j-1}}{365^{j-1}} \end{aligned}$$

And if $X \geq j$, then $I_j = 1$, if $X \leq j$, then I_j=0, we can get $P(X \geq j) = P(I_j = 1) = E(I_j)$ so we can prove that $X = \sum_{j=1}^{366}$, and we can find $E(X)$ by

$$E(X) = E(\sum_{j=1}^{366} I_j) = \sum_{j=1}^{366} E(I_j) = \sum_{j=1}^{366} p_j$$

(c)

Calculate the result in computer, we can get $E(X) \approx 24.62$

(d)

According to the properties of indicator, we have $I_i^2 = I_i$ and $I_iI_j = I_j$, thus we can conclude that

$$\begin{aligned} X_2 &= I_1^2 + ... + I_{366}^2 + 2\sum_{j=2}^{366}\sum_{i=1}^{j-1} I_iI_j \\ &= I_1 + ... + I_{366} + 2\sum_{j=2}^{366}(j-1)I_j \end{aligned}$$

$$E(X^2) = p_1 + ... + p_{366} + 2\sum_{j=2}^{366}(j-1)p_j = 2\sum_{j=2}^{366}(j-1)p_j$$

$$Var(X) = E(X^2) - (E(X))^2 = 2\sum_{j=2}^{366}(j-1)p_j - (\sum_{j=1}^{366} p_j)^2 \approx 148.64$$

Problem 6

1. **Step 1**

$$p_0 = 0 \quad p_1 = 0 \quad p_2 = 0 \quad p_3 = (\frac{1}{2})^3 = \frac{1}{8} \quad p^4 = (1 - \frac{1}{2}) \times (\frac{1}{2})^3 = \frac{1}{16}$$

Step 2

when $k \geq 4$, we denote that S_1:result of the first toss, then we have

$$\begin{aligned}
p_k &= P(N = k) \\
&= P(N = k, S_1 = H) + P(N = k, S_1 = T) \\
P(N = k, S_1 = H) &= P(S_1 = H)P(S_2 = H)P(N = k - 2) + P(S_1 = H)P(S_2 = T)P(N = k - 2) \\
&= P(S_1 = H)P(S_2 = H)P(S_3 = T)P(N = k - 3) + P(S_1 = H)P(S_2 = T)P(N = \\
&= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times p_{k-3} + \frac{1}{2} \times \frac{1}{2} \times p_{k-2} \\
&= \frac{1}{8}p_{k-3} + \frac{1}{4}p_{k-2} \\
P(N = k, S_1 = T) &= P(S_1 = T)P(N = k - 1) \\
&= \frac{1}{2} \times p_{k-1} \\
&= \frac{1}{2}p_{k-1}
\end{aligned}$$

Thus,

$$p_k = \frac{1}{8}p_{k-3} + \frac{1}{4}p_{k-2} + \frac{1}{2}p_{k-1}$$

then we can get the probability is

$$p_k = \begin{cases} 0 & k = 0 \\ 0 & k = 1 \\ 0 & k = 2 \\ \frac{1}{8} & k = 3 \\ \frac{1}{8}p_{k-3} + \frac{1}{4}p_{k-2} + \frac{1}{2}p_{k-1} & k \geq 4 \end{cases}$$

Step 3

PGF of N, on one hand

$$g(t) = E(t^N) = \sum_{k=0}^{\infty} t^k = \sum_{k=0}^{\infty} p_k t^k = \sum_{k=1}^{\infty} p_k t^k = \sum_{k=2}^{\infty} p_k t^k = \sum_{k=3}^{\infty} p_k t^k = p_3 t^3 + \sum_{k=4}^{\infty} p_k t^k = \frac{1}{8}t^3 + \sum_{k=4}^{\infty} p_k t^k$$

On the other hand

$$\begin{aligned}
p_k &= \frac{1}{8}p_{k-3} + \frac{1}{4}p_{k-2} + \frac{1}{2}p_{k-1} \qquad (k \geq 4) \\
\sum_{k=4}^{\infty} p_k t^k &= \sum_{k=4}^{\infty}(\frac{1}{8}p_{k-3} + \frac{1}{4}p_{k-2} + \frac{1}{2}p_{k-1})t^k \\
&= \frac{1}{8}\sum_{k=4}^{\infty} p_{k-3}t^k + \frac{1}{4}\sum_{k=4}^{\infty} p_{k-2}t^k + \frac{1}{2}\sum_{k=4}^{\infty} p_{k-1}t^k \\
&= \frac{1}{8}t^3\sum_{k=4}^{\infty} p_{k-3}t^{k-3} + \frac{1}{4}t^2\sum_{k=4}^{\infty} p_{k-2}t^{k-2} + \frac{1}{2}t\sum_{k=4}^{\infty} p_{k-1}t^{k-1} \\
&= \frac{1}{8}t^3\sum_{k=1}^{\infty} p_k t^k + \frac{1}{4}t^2\sum_{k=2}^{\infty} p_k t^k + \frac{1}{2}t\sum_{k=3}^{\infty} p_k t^k \\
&= (\frac{1}{8}t^3 + \frac{1}{4}t^2 + \frac{1}{2}t)g(t) = g(t) - \frac{1}{8}t^3
\end{aligned}$$

Step 4

Thus, we get

$$\begin{aligned}
g(t) &= \frac{\frac{1}{8}t^3}{1 - (\frac{1}{8}t^3 + \frac{1}{4}t^2 + \frac{1}{2}t)} \\
E(N) &= g'(t)|_{t=1} = g'(1) = 14 \\
Var(N) &= g''(1) + g'(1) - [g'(1)]^2 = 142
\end{aligned}$$