# Probability & Statistics for EECS: Homework #11

Due on Jan 1, 2024 at 23:59

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

# Problem 1

1. (a)

If $X \sim \text{Geom}(p)$, then$P(X = k) = q^k p$ where $q = 1 - p$.
which means that $p_k = P(X = x_k) = pq^k$

$$\begin{aligned}
H(X) &= \sum_k P(X = k) \log_2 (\frac{1}{P(X = k)}) \\
&= \sum_k p_k \log_2 (p_k^{-1}) \\
&= -\sum_k q^k p \log_2 (q^k p) \\
&= -\log_2 p \sum_k q^k p - \log_2 (q) \sum_k k q^k p \\
&= -\log_2 p \cdot 1 - \log_2 (q) \cdot E(X) \\
&= -\log_2 p - \frac{1-p}{p} \log_2 (1 - p)
\end{aligned}$$

(b)

Let $Z$ be an r.v. whose PMF is $P(Z = p_k) = p_k$.

$$\begin{aligned}
E(\log_2 Z) &= \sum_k p_k \log_2 (p_k) \\
&= -\sum_k p_k \log_2 (\frac{1}{p_k}) \\
&= -H(X)
\end{aligned}$$

$$\begin{aligned}
\log_2 E(Z) &= \log_2 \left( \sum_k p_k^2 \right) \\
&= \log_2 \left( \sum_k P(X = k) \cdot P(Y = k) \right) \\
&= \log_2 P(X = Y)
\end{aligned}$$

Using Jensen's inequality,

$$E(\log_2 Z) \le \log_2 E(Z)$$

Then we get that $-H(X) \ge \log_2 P(X = Y)$ which is

$$P(X = Y) \ge 2^{-H(X)}$$

# Problem 2

1. (a)

   $X \sim Pois(\lambda)$ gives $P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

$$\begin{aligned} P(X=k|X\geq 1) &= \frac{P(X=k)}{1-P(X=0)} \\ &= \frac{\frac{e^{-\lambda}\lambda^k}{k!}}{1-e^{-\lambda}} \\ &= \frac{e^{-\lambda}\lambda^k}{(1-e^{-\lambda})k!}, \quad k \geq 1 \end{aligned}$$

$$\begin{aligned} E[X|X\geq 1] &= \sum_{k=1}^{\infty} k \cdot P(X=k|X\geq 1) \\ &= \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{k\lambda^k}{k!} \\ &= \frac{\lambda e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \frac{\lambda e^{-\lambda}}{1-e^{-\lambda}} \cdot e^{\lambda} \\ &= \frac{\lambda}{1-e^{-\lambda}} \end{aligned}$$

   (b)

$$\begin{aligned} E[X^2|X\geq 1] &= \sum_{k=1}^{\infty} k^2 \cdot P(X=k|X\geq 1) \\ &= \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{k^2\lambda^k}{k!} \\ &= \frac{\lambda e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{k\lambda^{k-1}}{(k-1)!} \\ &= \frac{\lambda e^{-\lambda}}{1-e^{-\lambda}} (e^{\lambda} + \lambda e^{\lambda}) \\ &= \frac{\lambda(\lambda+1)}{1-e^{-\lambda}} \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X|X\geq 1) &= E[X^2|X\geq 1] - (E[X|X\geq 1])^2 \\ &= \frac{\lambda(\lambda+1)}{1-e^{-\lambda}} - \left(\frac{\lambda}{1-e^{-\lambda}}\right)^2 \\ &= \frac{\lambda(\lambda+1)(1-e^{-\lambda}) - \lambda^2}{(1-e^{-\lambda})^2} \\ &= \frac{\lambda - \lambda e^{-\lambda} - \lambda^2 e^{-\lambda}}{(1-e^{-\lambda})^2} \\ &= \frac{\lambda(1-e^{-\lambda}(1+\lambda))}{(1-e^{-\lambda})^2} \end{aligned}$$

# Problem 3

1. (a)

   The memoryless property of an exponential distribution means

$$P(X_1 > s + t | X_1 > s) = P(X_1 > t)$$

   Therefore, the conditional expectation $E[X_1 | X_1 > 2023]$ is:

$$\begin{aligned} E[X_1 | X_1 > 2023] &= E[X_1 - 2023 + 2023 | X_1 > 2023] \\ &= E[X_1 - 2023 | X_1 > 2023] + 2023 \\ &= E[X_1] + 2023 \\ &= 2023 + \frac{1}{\lambda_1} \end{aligned}$$

   (b)

   Using the memoryless property of the exponential distribution for each $X_i$, we have:

$$\begin{aligned} E[X_1 + X_2 + X_3 | X_1 > 2023, X_2 > 2024, X_3 > 2025] &= E[X_1 | X_1 > 2023] + E[X_2 | X_2 > 2024] + E[X_3 | X_3 > 2025] \\ &= (2023 + \frac{1}{\lambda_1}) + (2024 + \frac{1}{\lambda_2}) + (2025 + \frac{1}{\lambda_3}) \\ &= 6072 + \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} \end{aligned}$$

# Problem 4

1. (a)

   Integrating the joint PDF over $x$:

$$f_Y(y) = \int_{y^2}^{1} 6xy\,dx = 3y[x^2]_0^1 = 3y(1-y^4)$$

   For $y \in [0,1]$ and 0 otherwise.
   Integrating the joint PDF over $y$:

$$f_X(x) = \int_{0}^{\sqrt{x}} 6xy\,dy = 3x[y^2]_0^{\sqrt{x}} = 3x^2$$

   For $x \in [0,1]$ and 0 otherwise.
   And we can find that

$$f_X(x)f_Y(y) = (3y)(3x^2) = 9x^2y \neq 6xy = f_{X,Y}(x,y)$$

   Thus, $X$ and $Y$ are not independent.

   (b)

$$E[X|Y=y] = \int_{y^2}^{1} x \cdot \frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx = \int_{0}^{1} x \cdot \frac{6xy}{3y(1-y^4)}\,dx = \frac{2(1-y^6)}{3(1-y^4)}$$

$$E[X^2|Y=y] = \int_{y^2}^{1} x^2 \cdot \frac{6xy}{3y(1-y^4)}\,dx = \frac{1+y^4}{2}$$

   the variance is:

$$Var[X|Y=y] = E[X^2|Y=y] - (E[X|Y=y])^2 = \frac{1+y^4}{2} - \left(\frac{2(1-y^6)}{3(1-y^4)}\right)^2 = \frac{y^{12}-9y^8+16y^6-9y^4+1}{18(1-y^4)^2}$$

   (c)

   From (b) we cn find that

$$E[X|Y] = \frac{2(1-Y^6)}{3(1-Y^4)}$$

$$Var[X|Y] = \frac{Y^{12}-9Y^8+16Y^6-9Y^4+1}{18(1-Y^4)^2}$$

# Problem 5

1. (a)

   the sample mean of the Bernoulli trials: $\hat{p} = \frac{1}{N}\sum_{i=1}^{N} X_i$. The variance of a Bernoulli trial is $p(1-p)$, and thus the variance of the sample mean $\hat{p}$ is $\frac{p(1-p)}{N}$. We want to find the value of $\varepsilon$ such that:

   $$P(|\hat{p} - p| \geq \varepsilon) \leq \delta$$

   Using Chebyshev's inequality, we have:

   $$\delta \geq \frac{\sigma^2}{\varepsilon^2}$$

   which means that

   $$\delta \geq \frac{p(1-p)}{N\varepsilon^2}$$

   Solve $\varepsilon$ to find the confidence interval $p \pm \varepsilon$. So we get:$\varepsilon \geq \sqrt{\frac{p(1-p)}{N\delta}}$
   Thus the confidence interval is

   $$p \in (\hat{p} - \epsilon, \hat{p} + \epsilon)$$

   where

   $$\epsilon = \sqrt{\frac{\hat{p}(1-\hat{p})}{N\delta}}$$

   (b)

   The Hoeffding bound for the sum of $N$ i.i.d. random variables $X_i$ is:

   $$P\left(\left|\frac{1}{N}\sum_{i=1}^{N} X_i - \mu\right| \geq \varepsilon\right) \leq 2\exp\left(-2N\varepsilon^2\right)$$

   Setting $2\exp\left(-2N\varepsilon^2\right) = \delta$ and solving for $\varepsilon$ gives:

   $$\varepsilon = \sqrt{\frac{\ln(2/\delta)}{2N}}$$

   . The confidence interval is thus:

   $$p \in (\hat{p} - \varepsilon, \hat{p} + \varepsilon)$$

   where

   $$\varepsilon = \sqrt{\frac{\ln(2/\delta)}{2N}}$$

   discuss the Impact of $\delta$ and $N$

   For both methods, as $\delta$ decreases, the confidence level $1-\delta$ increases, leading to a wider interval. As $N$ increases, the interval narrows, indicating a more precise estimate of $p$.

   (d)

   The Chebyshev inequality does not require the underlying distribution to be known, but it often provides a looser bound than the Hoeffding bound. The Hoeffding bound gives a tighter interval but assumes that the random variables are bounded and i.i.d.

# Problem 6

1. (a)

$$\frac{1}{p} + \frac{1}{1-p}$$

(b)

Suppose that we need $X$ times for $HH$

$$\begin{aligned} E(X) &= E(X|H)p + E(X|T)(1-p) \\ &= E(X|H)p + (1+E(X))(1-p) \end{aligned}$$

On the other hand,

$$\begin{aligned} E(X|H) &= E(X|HH)p + E(X|HT)(1-p) \\ &= 2p + (2+E(X))(1-p). \end{aligned}$$

Thus, we can get

$$\begin{aligned} E(X) &= E(X|H)p + (1+E(X))(1-p) \\ &= (2p + (2+E(X))(1-p))p + (1+E(X))(1-p) \end{aligned}$$

Simplify this to get $E(X) = \frac{1}{p} + \frac{1}{p^2}$

(c)