E1 222 Stochastic Models and Applications
Test I

`Time: 75 minutes` `Max. Marks:40`
`Date: 23 Nov 2020`

Answer **ALL** questions. All questions carry equal marks

1. a. Let A, B, C be events in a probability space. If B and C are independent show that

$$P(A|B) = P(A|BC)P(C) + P(A|BC^c)P(C^c)$$

Answer: Since B and C are independent, so are B and C^c and we have $P(BC) = P(B)P(C)$ and $P(BC^c) = P(B)P(C^c)$. Hence we get

$$\begin{aligned} P(A|BC)P(C) + P(A|BC^c)P(C^c) &= \frac{P(ABC)}{P(BC)}P(C) + \frac{P(ABC^c)}{P(BC^c)}P(C^c) \\ &= \frac{P(ABC)}{P(B)} + \frac{P(ABC^c)}{P(B)} \\ &= \frac{P(ABC) + P(ABC^c)}{P(B)} \\ &= \frac{P(AB)}{P(B)} = P(A|B) \end{aligned}$$

b. Two numbers are drawn at random with replacement from $\{1, 2, \cdots, N\}$. Calculate the probability that one of the numbers is less than or equal to half of the other number. Assume N is even.

Answer : I will consider both cases of N being even and N being odd, though you need to consider only the case of N being even.

Suppose the first number is k, then the second number can be any of the numbers from $2k$ to N. Thus probabilty that the first number is k and the second number is at least twice the second number is given by $\frac{1}{N} \frac{N-2k+1}{N}$ provided $N - 2k + 1 \geq 1$. There would be at least one choice for the second number only if $2k \leq N$ or $k \leq \frac{N}{2}$.

Let A be the event that the first number is less than or equal to half of the second number. Assume N is even. Then, from the above, we have

$$\begin{aligned} P(A) &= \sum_{k=1}^{N/2} \frac{1}{N} \frac{N-2k+1}{N} \\ &= \frac{1}{N^2} \left(\frac{N}{2} N - 2\frac{N}{2} \left(\frac{N}{2} + 1 \right) \frac{1}{2} + \frac{N}{2} \right) \\ &= \frac{1}{4} \end{aligned}$$

On the otherhand, if N is odd, the highest value k can take in the above summation is $(N-1)/2$. Hence we get

$$\begin{aligned} P(A) &= \frac{1}{N^2} \left(\frac{N-1}{2} N - 2\frac{N-1}{2} \left(\frac{N-1}{2} + 1 \right) \frac{1}{2} + \frac{N-1}{2} \right) \\ &= \frac{1}{4} - \frac{1}{4N^2} \end{aligned}$$

The probability we want is $2P(A)$ because we can have either the second number at least twice the first number or the first number being at least twice the second number.

Hence the required probability is 0.5 when N is even and $0.5(1 - \frac{1}{N^2})$ when N is odd.

Q1. part(a) has 4 marks and part(b) has 6 marks

2. a. Let X be a continuous random variable with density function

$$f_X(x) = K(3x - x^2), \quad 0 < x < 2$$

Find the value of K, F_X, EX, and $P[X < 1]$.

Answer: We have

$$\int_0^2 K(3x - x^2)\, dx = K \left(3\frac{x^2}{2}\bigg|_0^2 - \frac{x^3}{3}\bigg|_0^2 \right) = K\, \frac{10}{3}$$

which implies $K = 0.3$.

The distribution function is given by, for $0 \le x \le 2$,

$$F_X(x) = \frac{3}{10} \int_0^x (3x - x^2)\, dx = \frac{3}{10} \left(3\frac{x^2}{2}\bigg|_0^x - \frac{x^3}{3}\bigg|_0^x \right) = \frac{1}{20}(9x^2 - 2x^3)$$

For $x < 0$, $F_X(x) = 0$ and for $x > 2$, $F_X(x) = 1$. We calculate EX as

$$EX = \frac{3}{10}\int_0^2 x(3x - x^2)\, dx = \frac{3}{10}\left(3\frac{x^3}{3}\bigg|_0^2 - \frac{x^4}{4}\bigg|_0^2\right) = \frac{12}{10}$$

Since X is a continuous random variable, $P[X < 1] = P[X \le 1] = F_X(1) = \frac{7}{20}$.

b. Suppose X is a continuous random variable with density $f_X(x) = -\ln(x),\ 0 < x < 1$. Find its distribution function. Let $Y = X - X\ln(X)$. Find the density of Y and EY.

Answer: The distribution function of X, for $0 < x < 1$, is given by

$$F_X(x) \;=\; \int_0^x (-\ln(x))\, dx = -\,x\ln(x)|_0^x + \int_0^x dx = x - x\ln(x)$$

Hence the distribution function is

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ x - x\ln(x) & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

It is given that $Y = X - X\ln(X)$. So, $Y = F_X(X)$. Hence we know that Y would be uniform over $[0,1]$ and hence $EY = 0.5$. We can find distribution of Y as follows.

We note that $\frac{d}{dx}(x - x\ln(x)) = 1 - \ln(x) - 1 = -\ln(x) > 0$ for $0 < x < 1$. This implies that F_X is strictly monotone and hence is invertible. Let us denote the inverse by F_X^{-1}. Since $Y = F_X(X)$ we know Y takes values in $[0,1]$. For $0 \le y \le 1$, we have

$$\begin{aligned} F_Y(y) &= P[Y \le y] \\ &= P[F_X(X) \le y] \\ &= P[X \le F_X^{-1}(y)], \quad \text{because } F_X \text{ is monotone increasing} \\ &= F_X\left(F_X^{-1}(y)\right) \\ &= y \end{aligned}$$

This shows that Y is uniform over $[0,1]$. Hence, $EY = 0.5$.

Q2. part(a) has 4 marks and part(b) has 6 marks

3. a. Let X be a continuous random variable having exponential density with parameter λ. For any given $\epsilon > 0$, let X_ϵ be defined by

$$X_\epsilon = \epsilon k \text{ if } \epsilon k \leq X < \epsilon(k+1), \ k \text{ integer.}$$

Find EX_ϵ and its limit as $\epsilon \to 0$.

Answer: Given a specific $\epsilon > 0$, the possible values of X_ϵ are: $0, \epsilon, 2\epsilon, \cdots$. It is a discrete random variable and its pmf is given by

$$\begin{aligned}
f_{X_\epsilon}(k\epsilon) &= P[X_\epsilon = k\epsilon] \\
&= P[\epsilon k \leq X < \epsilon(k+1)] \\
&= F_X(\epsilon(k+1)) - F_X(\epsilon k) \\
&= e^{-\lambda\epsilon k} - e^{-\lambda\epsilon(k+1)}, \quad \text{because } X \text{ is exponential} \\
&= e^{-\lambda\epsilon k}\left(1 - e^{-\lambda\epsilon}\right)
\end{aligned}$$

Now we can calculate its expectation as

$$\begin{aligned}
E[X_\epsilon] &= \sum_{k=0}^{\infty} k\epsilon \, e^{-\lambda\epsilon k}\left(1 - e^{-\lambda\epsilon}\right) \\
&= \left(1 - e^{-\lambda\epsilon}\right)\sum_{k=0}^{\infty} k\epsilon \, e^{-\lambda\epsilon k}
\end{aligned}$$

We know that

$$\sum_{k=0}^{\infty} e^{-\lambda\epsilon k} = \frac{1}{(1 - e^{-\lambda\epsilon})}$$

Differentiating both sides with respect to λ, we get

$$\sum_{k=0}^{\infty} k\epsilon \, e^{-\lambda\epsilon k} = \frac{\epsilon \, e^{-\lambda\epsilon}}{(1 - e^{-\lambda\epsilon})^2}$$

Thus we get

$$EX_\epsilon = \left(1 - e^{-\lambda\epsilon}\right)\frac{\epsilon \, e^{-\lambda\epsilon}}{(1 - e^{-\lambda\epsilon})^2} = \frac{\epsilon \, e^{-\lambda\epsilon}}{(1 - e^{-\lambda\epsilon})}$$

We need to find the limit of the above as $\epsilon \to 0$. If we put $\epsilon = 0$ in the above expression, we get the indeterminate form of 0/0. Hence using L'Hospital's rule

$$\lim_{\epsilon \to 0} E[X_\epsilon] = \lim_{\epsilon \to 0} \frac{e^{-\lambda\epsilon} - \epsilon\lambda e^{-\lambda\epsilon}}{\lambda e^{-\lambda\epsilon}} = \frac{1}{\lambda}$$

b. Let X be a discrete random variable having geometric distribution with parameter p. Let $M > 0$ be an integer. Define $Y = \max(X, M)$. Find distribution of Y.

Answer: Since $Y = \max(X, M)$, we know $Y \geq M$. If $X \leq M$ then $Y = M$; else $Y = X$. Thus the pmf of Y is given by

$$f_Y(k) = P[Y = k] = \begin{cases} 0 & \text{if } \; k < M \\ P[X \leq M] = 1 - (1-p)^M & \text{if } \; k = M \\ P[X = k] = (1-p)^{k-1}p & \text{if } \; k > M \end{cases}$$

Q3. part(a) has 6 marks and part(b) has 4 marks

4. a. Let X be a non-negative integer valued random variable. Let $\Phi_X(t) = Et^X$ be its probability generating function and assume that $\Phi_X(t)$ is finite for all t. Show that for any positive integer, y,

$$P[X \leq y] \leq \frac{\Phi_X(t)}{t^y}, \; 0 \leq t \leq 1$$

Answer: We have, for $0 < t < 1$,

$$\begin{aligned} \Phi_X(t) &= E\left[t^X\right] = \sum_{k=0}^{\infty} t^k \, f_X(k) \\ &= \sum_{k=0}^{y} t^k \, f_X(k) \; + \sum_{k=y+1}^{\infty} t^k \, f_X(k) \\ &\geq \sum_{k=0}^{y} t^k \, f_X(k), \quad \text{since all terms in the summation are non-negative} \\ &\geq t^y \sum_{k=0}^{y} f_X(k), \quad \text{because, for } 0 < t < 1, \, t^y \leq t^k \text{ for } k \leq y \\ &= t^y P[X \leq y] \end{aligned}$$

This gives us the required relation:

$$P[X \leq y] \leq \frac{\Phi_X(t)}{t^y}$$

Using Φ_X we can also bound $P[X \geq y]$. I give that also below because it is a useful bound. This is obviously not needed in the test because it is not part of the question.

For this we consider the case $t \geq 1$

$$\begin{aligned}
\Phi_X(t) &= E\left[t^X\right] = \sum_{k=0}^{\infty} t^k\, f_X(k) \\
&= \sum_{k=0}^{y-1} t^k\, f_X(k) \;+\; \sum_{k=y}^{\infty} t^k\, f_X(k) \\
&\geq \sum_{k=y}^{\infty} t^k\, f_X(k), \quad \text{since all terms in the summation are non-negative} \\
&\geq t^y \sum_{k=y}^{\infty} f_X(k), \quad \text{because, for } t > 1,\ t^y \leq t^k \text{ for } k \geq y \\
&= t^y P[X \geq y]
\end{aligned}$$

This gives us the relation

$$P[X \geq y] \leq \frac{\Phi_X(t)}{t^y},\ t \geq 1$$

b. Let X, Y be random variables with joint density given by

$$f_{XY}(x,y) = e^{-y},\ \ 0 < x < y < \infty$$

Find the marginal densities of X, Y and $P[X \leq \frac{y}{2}|Y = y]$

Answer: The marginals can be calculated as follows

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy = \int_x^{\infty} e^{-y}\, dy = e^{-x},\ \ x > 0$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx = \int_0^{y} e^{-y}\, dx = y\, e^{-y}, y > 0$$

Thus, X is exponential and Y is gamma.

The conditional density $f_{X|Y}$ is

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y},\ \ 0 < x < y < \infty$$

Hence the conditional density of X given Y is uniform from 0 to Y. Hence, $P[X \leq \frac{y}{2}|Y = y] = 0.5$. We can explicitly compute it as

$$P\left[X \leq \frac{y}{2}|Y = y\right] = \int_{-\infty}^{\frac{y}{2}} f_{X|Y}(x|y)\, dx = \int_0^{\frac{y}{2}} \frac{1}{y}\, dx = \frac{1}{2}$$

Q4. part(a) has 5 marks and part(b) has 5 marks