E1 222 Stochastic Models and Applications
Test I

Time: 90 minutes | Max. Marks:40
Date: 21 Sept 2019

Answer **ALL** questions. All questions carry equal marks
Answers should be written only in the space provided.

1. a. A coin, whose probability of heads is p, is tossed repeatedly till we get a head. Calculate the probability that the number of tosses needed is odd.

Answer: Probability that it takes k tosses is given by $(1-p)^{k-1}p$, $k = 1, 2, \cdots$. We need to sum this over all odd k to get the required probability. We take $k = 2m+1$ and sum over $m = 0, 1, \cdots$, to sum over all odd k. Hence the required probability is

$$\sum_{m=0}^{\infty}(1-p)^{2m}p = \frac{p}{1-(1-p)^2} = \frac{1}{2-p}$$

b. A point S is chosen at random on a line segment AB. Show that the probability that the length of AS divided by the length of SB is smaller than a is $\frac{a}{1+a}$ (where we take $a > 0$). Also show that the probability that the length of the longer segment is greater than three times the length of the shorter segment is 0.5.

Answer: Let $|AS|$ denote length of line segment AS and similarly for others. Then we have

$$\frac{|AS|}{|SB|} < a \;\Rightarrow\; \frac{|AS|}{|AB|-|AS|} < a \;\Rightarrow\; \frac{|AS|}{|AB|} < \frac{a}{1+a}$$

Since the point S is uniformly distributed over line segment AB, the above shows that the probability of $\frac{|AS|}{|SB|} < a$ is $\frac{a}{1+a}$.
Let x denote the length of one segment. Then $L-x$ would be the length of the other segment (assuming L is length of AB). So we need either $x > 3(L-x)$, (which is same as $x > \frac{3}{4}L$) or $(L-x) > 3x$ (which is same as $x < \frac{1}{4}L$). Hence the required probability is $0.25+0.25 = 0.5$.

We can also do this more formally using the language of random variables. Let length of AB be L. Let the length of AS be denoted by random variable X. Then S being chosen randomly on line segment AB means X is uniform over $[0,\ L]$. Now for the first part we have

$$P\left[\frac{X}{L-X} < a\right] = P\left[X < \frac{a}{1+a}L\right] = \left(\frac{a}{1+a}L\right)/L = \frac{a}{1+a}$$

For the second part we have

$$\begin{aligned}
P[\max(X, L-X) > 3\min(X, L-X)] &= P\left[X < \frac{L}{2}, \max(X, L-X) > 3\min(X, L-X)\right] \\
&\quad + P\left[X \geq \frac{L}{2}, \max(X, L-X) > 3\min(X, L-X)\right] \\
&= P\left[X < \frac{L}{2}, L-X > 3X\right] + P\left[X \geq \frac{L}{2}, X > 3(L-X)\right] \\
&= P\left[X < \frac{L}{2}, X < \frac{L}{4}\right] + P\left[X \geq \frac{L}{2}, X > \frac{3L}{4}\right] \\
&= 0.25 + 0.25 = 0.5
\end{aligned}$$

(Here we have used the obvious facts that when $X < \frac{L}{2}$, we have $\max(X, L-X) = L - X$ and so on).

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

$$f_X(x) = K(1-x)^3,\ 0 \leq x \leq 1$$

Let $Y = X^2$. Find $P[X > 0.5]$, $F_X(x)$, EY and variance of X.

Answer: Since f_X has to be a density, we need $\int_0^1 K(1-x)^3\, dx = 1$ which gives $K = 4$. Since density is zero outside $[0,\ 1]$, we have $F_X(x) = 0$ for $x < 0$ and $F_X(x) = 1$ for $x > 1$. For, $0 \leq x \leq 1$, we have

$$F_x(x) = \int_0^x 4(1-x)^3\, dx = 1 - (1-x)^4$$

Hence, $P[X > 0.5] = 1 - F_X(0.5) = (1-0.5)^4 = 1/16$.
We have $EX = \int_0^1 4x(1-x)^3\, dx$. We can solve this using integration by parts or by writing it as $\int_0^1 4x(1 - x^3 - 3x + 3x^2)\, dx$. This gives us $EX = \frac{1}{5}$.

We have $EX^2 = \int_0^1 4x^2(1-x)^3\ dx = \int_0^1 4x^2(1-x^3-3x+3x^2)\ dx$. We get $EX^2 = \frac{1}{15}$.

Hence $EY = EX^2 = \frac{1}{15}$ and $\text{Var}(X) = EX^2 - (EX)^2 = \frac{1}{15} - \frac{1}{25} = \frac{2}{75}$.

b. Consider a random experiment with $\Omega = \{(x,y) \in \Re^2 : -1 \leq x, y \leq +1\}$. Note that each outcome of this experiment is a point in a square of side 2 centered at the origin. This experiment is repeated till the outcome or the point obtained is such that the distance of the point from the origin is less than or equal to one. Let the random variable X denote the number of repetitions needed. What is the probability mass function of X?

Answer: The distance of the point from origin being less than or equal to 1 is same as the point being in or on the circle with center at origin and radius 1. This circle is completely in the unit square. Its area is π. The area of the square is 4. Hence the probability of the point falling in the circle on any repetition of the experiment is $\pi/4$. We would need k repetitions if for the first $k-1$ times the point does not fall in the circle and the k^{th} time it does. Hence we have $f_X(k) = (1-\pi/4)^{k-1}(\pi/4)$, $k = 1, 2, \cdots$.

Comment: Suppose (X,Y) is a two dimensional random vector whose density is constant over the unit disc (that is the region in and on the unit circle) and is zero outside. That is, (X,Y) is uniform over the unit disc. Suppose we want to generate such a two dimensional random vector uniformly distributed over the unit disc. We have access only to a random variable that is uniform over $[0,\ 1]$. Does this problem suggest a method of generating (X,Y) uniform over the unit disc?

3. a. Let X be a continuous random variable that is uniform over $[a,\ b]$. Let $Y = (X-c)^2$ where $\frac{a+b}{2} < c < b$. Find the probability density function of Y.

Answer: The density of X is given as

$$f_x(x) = \frac{1}{(b-a)}, \quad a \leq x \leq b$$

The given condition of c implies that c is closer to b than to a and hence we have $(a-c)^2 > (b-c)^2$. Hence the range of Y is 0 to $(a-c)^2$. For $y \geq 0$, we have

$$\begin{aligned} F_Y(y) &= P[(X-c)^2 \leq y] \\ &= P[-\sqrt{y} \leq (X-c) \leq \sqrt{y}] \\ &= P[c-\sqrt{y} \leq X \leq c+\sqrt{y}] \\ &= F_X(c+\sqrt{y}) - F_X(c-\sqrt{y}) \end{aligned}$$

By differentiating this, we get density of Y as

$$f_Y(y) = \frac{1}{2\sqrt{y}} \left(f_X(c+\sqrt{y}) + f_X(c-\sqrt{y})\right)$$

Given the condition on c, for $y > (a-c)^2$, $c-\sqrt{y} < a$ and $c+\sqrt{y} > b$. So, $f_X(c+\sqrt{y}) = f_X(c-\sqrt{y}) = 0$ and hence the density of Y is zero in this range, as we already know. Since c is closer to b than a, when $(b-c)^2 \leq y \leq (a-c)^2$, $c+\sqrt{y} > b$ and hence $f_X(c+\sqrt{y}) = 0$. Putting all this together, we can now write density of Y as

$$\begin{aligned} f_Y(y) &= \frac{1}{2\sqrt{y}} \left(\frac{1}{b-a} + \frac{1}{b-a}\right) = \frac{1}{\sqrt{y}(b-a)}, \quad 0 < y < (b-c)^2 \\ &= \frac{1}{2\sqrt{y}} \left(0 + \frac{1}{b-a}\right) = \frac{1}{2\sqrt{y}(b-a)}, \quad (b-c)^2 \leq y \leq (a-c)^2 \end{aligned}$$

Comment: Could we have used the formula given in class for finding density of a function of a continuous random variable? The answer is 'No' because the given function is not monotone. (If you take $h(x) = (x-c)^2$ then $h'(x) = 2(x-c)$ and since x takes values that may be greater than or less than c, the derivative can be both positive and negative which means h is not monotone). We know that squaring is not an invertible function because it is not one-to-one. (E.g., both $+1$ and -1 map to $+1$). In class we showed that if $Y = X^2$ (and X is continuous) then $f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y})+f_X(-\sqrt{y}))$, $y > 0$. This is essentially same as what we derived above. If we wrongly thought $h(x) = x^2$ is invertible with $h^{-1}(y) = \sqrt{y}$ then we would derive $f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y})$ which is wrong.

b. Let X be a continuous random variable with density: $f_X(x) = 3e^{-3x}, \; x > 0$. Let $Y = 1 - e^{-3X}$. Find the density of Y.

Answer: A simple calculation (or recall from memory of the exponential distribution function) gives us $F_X(x) = 1 - e^{-3x}, \; x > 0$. Since density of X is zero for negative values, we know $X \geq 0$. Hence $0 \leq e^{-3X} \leq 1$ and hence range of Y is 0 to 1. Take $y \in [0, \; 1]$. Now we have

$$
\begin{aligned}
F_Y(y) &= P[1 - e^{-3X} \leq y] \\
&= P[e^{-3X} \geq 1 - y] \\
&= P[-3X \geq \ln(1 - y)] \\
&= P[X \leq -\frac{1}{3} \ln(1 - y)] \\
&= F_X\left(\frac{-1}{3} \ln(1 - y)\right) \\
&= 1 - e^{-3(\frac{-1}{3} \ln(1-y))} = 1 - (1 - y) = y
\end{aligned}
$$

This shows that Y is uniform over $[0, \; 1]$.

Comment: Suppose F is a continuous strictly monotonically increasing distribution function. In the class we showed the following: if X is uniform over $[0, \; 1]$ and $Y = F^{-1}(X)$ then Y has distribution F. As we saw, this has applications, e.g., for random number generation. This problem shows the following: if X has distribution F and $Y = F(X)$ then Y is uniform over $[0, \; 1]$. This is also useful in applications. One well-known application of this is what is called histogram equalization in image processing.

In this problem we could have used the formula because the function is monotone. Take $h(x) = 1 - e^{-3x}$. This is monotone. (We know e^x is monotone. Even otherwise, we can see $h'(x) = 3e^{-3x}$ which is positive for all x). Simple algebra shows that $h^{-1}(y) = \frac{-1}{3} \ln(1 - y)$. Derivative of this is $\frac{1}{3} \frac{1}{1-y}$. Hence we get

$$f_Y(y) = 3e^{-3(\frac{-1}{3} \ln(1-y))} \frac{1}{3} \frac{1}{1 - y} = e^{\ln(1-y)} \frac{1}{1 - y} = 1$$

Note that our formula does not automatically tell us the range of y for which $f_Y(y)$ is given by the above. From the function h, we can see that the range is [0, 1] and hence that is the range over which density of Y can be nonzero.

4. a. Let X, Y be continuous random variables with joint density

$$f_{XY}(x, y) = 6y, \quad 0 < y < x < 1$$

Find the marginal densities and $P[X > 0.75 \mid Y = 0.5]$.

Answer: The marginal density for X is given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\ dy = \int_0^x 6y\ dy = 3x^2, \quad 0 < x < 1$$

The marginal density for Y is given by

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\ dx = \int_y^1 6y\ dx = 6y(1 - y), \quad 0 < y < 1$$

The conditional density of X given Y is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{6y}{6y(1 - y)} = \frac{1}{1 - y}, \quad 0 < y < x < 1$$

(Thus, conditioned on Y, X is uniform from Y to 1).
Now, we can get the conditional probability as

$$P[X > 0.75 \mid Y = 0.5] = \int_{0.75}^{1} f_{X|Y}(x|0.5)\ dx = \int_{0.75}^{1} \frac{1}{1 - 0.5}\ dx = 0.5$$

b. Suppose two students A and B are solving a problem. The times taken by the two students are independent random variables. The time taken by A is uniformly distributed in [1, 3] while that by B is uniformly distributed in [2, 4]. What is the probability that A takes longer than B to solve the problem.

Answer: For any two independent random variables, X, Y,

$$P[X > Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_X(x) f_Y(y)\ dy\ dx = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_Y(y)\ dy\ dx$$

We can take X to be time taken by A and Y to be time taken by B. Then, X is uniform over [1, 3]. Hence its density is zero outside [1, 3]. So, now we can write

$$P[X > Y] = \int_1^3 \frac{1}{2} \int_{-\infty}^{x} f_Y(y)\ dy\ dx$$

We are given that Y is uniform over $[2,\ 4]$. Hence, in the above, if $x < 2$ then in the y-integral, throughout the range, $f_Y(y)$ is zero. Hence we get

$$P[X > Y] = \int_2^3 \frac{1}{2} \int_2^x \frac{1}{2}\, dy\, dx = \frac{1}{8}$$

We can also solve this using geometry. Here (X, Y) is uniform over $[1,\ 3] \times [2,\ 4]$. So, actually we can think of this as Ω and take probability of any subset to be area of the subset divided by 4. The subset we want is $\{(x, y) \in \Omega : x > y\}$. It is easily seen to be a triangle with area 0.5.

Name:

S.R.Number:

E1 222 Stochastic Models and Applications
Test II

`Time: 90 minutes` `Max. Marks:40`
`Date: 26 Oct 2019`

Answer **ALL** questions. All questions carry equal marks
Answers should be written only in the space provided.

1. (a) Let X, Y have joint density given by

$$f_{XY}(x,y) = \lambda^2 e^{-\lambda y},\ 0 \le x \le y < \infty$$

Find marginal densities of X, Y, and $E[X|Y]$ and $\text{Cov}(X,Y)$.

Answer : For the marginal of X we have

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

(Note that we need not actually do the integration. We know the integral because we know the distribution function of exponential random variable).
Thus, X is exponential with parameter λ.
Similarly, for the marginal of Y

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\ dx = \int_0^{y} \lambda^2 e^{-\lambda y}\ dx = \lambda^2 y e^{-\lambda y}, 0 \le y < \infty$$

This is easily seen to be gamma density with parameters 2 and λ.
The conditional density of X given Y is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{\lambda^2 e^{-\lambda y}}{\lambda^2 y e^{-\lambda y}} = \frac{1}{y},\ 0 \le x \le y < \infty$$

This is a uniform density over $[0, y]$. That is, conditioned on Y, X is uniform over $[0, Y]$. Hence $E[X|Y] = \frac{Y}{2}$.

We can get this conditional expectation by direct calculation too.

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

Hence, $E[X|Y] = \frac{Y}{2}$

We know $\mathrm{Cov}(X,Y) = EXY - EXEY$. From the marginal densities, we see that X is exponential while Y is gamma. We have $EX = \frac{1}{\lambda}$ and $EY = \frac{2}{\lambda}$.

$$EXY = \int_0^{\infty}\int_0^y xy\ \lambda^2 e^{-\lambda y}\ dx\ dy = \int_0^{\infty} y\lambda^2 e^{-\lambda y}\frac{y^2}{2}dy = \frac{1}{2}\int_0^{\infty} y^2\ \lambda^2 y e^{-\lambda y}dy$$

We can see that the last integral above is the second moment of a Gamma random variable with parameters $\alpha = 2$ and λ. For any Z, $EZ^2 = \mathrm{Var}(Z) + (EZ)^2$. hence, we get

$$EXY = \frac{1}{2}\left(\frac{2}{\lambda^2} + \frac{4}{\lambda^2}\right) = \frac{3}{\lambda^2}$$

(Evaluating EXY by direct integration is also not too difficult). Now,

$$\mathrm{Cov}(X,Y) = \frac{3}{\lambda^2} - \frac{2}{\lambda}\frac{1}{\lambda} = \frac{1}{\lambda^2}$$

Comment: I hope at least some of you noticed that this is essentially an example problem that I solved in class. After defining conditional expectation, this problem was solved in class to illustrate how you calculate conditional expectation given the joint density. At that time also I remarked on how you can avoid doing integrations here by using known facts about distributions. In the class we used this joint density with $\lambda = 1$. Here you have a general λ. That is the only difference.

(b) Let X, Y be random variables with $EX = EY = 0$, $\mathrm{Var}(X) = 1$, $\mathrm{Var}(Y) = 2$. Let ρ be the correlation coefficiet of X, Y. Let $Z = X + Y$ and $W = X - Y$. Find $\mathrm{Var}(Z)$, $\mathrm{Var}(W)$ and $\mathrm{Cov}(Z, W)$.

Answer: We have $\mathrm{Cov}(X,Y) = \rho_{XY}\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)} = \rho\sqrt{1*2} = \rho\sqrt{2}$. Hence

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) = 3 + \rho\ 2\sqrt{2}$$

$$\mathrm{Var}(W) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X,Y) = 3 - \rho\, 2\sqrt{2}$$

$\mathrm{Cov}(Z,W) = EZW - EZEW$. Since $EX = EY = 0$, we have $EZ = EW = 0$.

$EZW = E[(X+Y)(X-Y)] = E[X^2]-E[Y^2] = \mathrm{Var}(X)-\mathrm{Var}(Y) = -1$

Hence, $\mathrm{Cov}(Z,W) = -1$.

2. (a) Let $X_1, X_2, \cdots, X_n$ be *iid* random variables each of them being uniform over $[0,\ 1]$. Let $Y_1 = X_1$, $Y_2 = X_1X_2$, $Y_3 = X_1X_2X_3$, and so on with $Y_n = X_1X_2\cdots, X_n$. Find the joint density of $Y_1, \cdots, Y_n$ and the conditional density of Y_3 given Y_1, Y_2.

Answer: The given transformation is invertible. The inverse transformation is given by

$$\begin{aligned}
X_1 &= Y_1 \\
X_2 &= Y_2/Y_1 \\
X_3 &= Y_3/Y_2 \\
&\cdot \\
&\cdot \\
&\cdot \\
X_n &= Y_n/Y_{n-1}
\end{aligned}$$

The Jacobian of the transformation is given by

$$\begin{vmatrix}
1 & 0 & 0 & \cdots & 0 \\
-y_2/(y_1^2) & 1/y_1 & 0 & \cdots & 0 \\
0 & -y_3/(y_2^2) & 1/y_2 & \cdots & 0 \\
 & & \cdots & & \\
0 & 0 & \cdots & -y_n/(y_{n-1}^2) & 1/y_{n-1}
\end{vmatrix} = \frac{1}{y_1y_2\cdots y_{n-1}}$$

The determinant is easy to evaluate because it is a triangular matrix. Now we get

$$f_{Y_1Y_2\cdots Y_n}(y_1, y_2, \cdots, y_n) = \left|\frac{1}{y_1y_2\cdots y_{n-1}}\right| f_{X_1X_2\cdots X_n}(y_1, y_2/y_1, y_3/y_2, \cdots, y_n/y_{n-1})$$

Since X_i are iid uniform over $[0,1]$, the joint density of $X_1, \cdots, X_n$ is zero unless all arguments are between 0 and 1. When all arguments are between 0 and 1, the joint density is 1. Hence, the joint density of $Y_1, Y_2, \cdots, Y_n$ is zero unless $0 \le y_1 \le 1$, $0 \le y_2/y_1 \le 1$ and so on till $0 \le y_n/y_{n-1} \le 1$. This gives us the joint density of $Y_1, Y_2, \cdots, Y_n$ as

$$f_{Y_1 Y_2 \cdots Y_n}(y_1, y_2, \cdots, y_n) = \frac{1}{y_1 y_2 \cdots y_{n-1}}, \ 0 \le y_n \le y_{n-1} \le \cdots \le y_2 \le y_1 \le 1$$

The above is true for all n. Hence we get

$$f_{Y_1 Y_2 Y_3}(y_1, y_2, y_3) = \frac{1}{y_1 y_2}, \ 0 \le y_3 \le y_2 \le y_1 \le 1$$

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{1}{y_1}, \ 0 \le y_2 \le y_1 \le 1$$

Hence the needed conditional density is

$$f_{Y3|Y_1 Y_2}(y_3|y_1, y_2) = \frac{f_{Y_1 Y_2 Y_3}(y_1, y_2, y_3)}{f_{Y_1 Y_2}(y_1, y_2)} = \frac{1}{y_2}, \ 0 \le y_3 \le y_2 \le y_1 \le 1$$

(b) Let X_1, X_2 be iid geometric random variables. Let $Y = \min(X_1, X_2)$. Show that Y is also geometric and find its expectation.

Answer: Since X_i are geometric, we have $P[X_i > y] = (1-p)^y$ for all positive integers y. Hence for any positive integer y,

$$P[Y > y] = P[\min(X_1, X_2) > y] = P[X_1 > y, X_2 > y]$$

$$= P[X_1 > y]P[X_2 > y] = (1-p)^y(1-p)^y = (1-p)^{2y}$$

So, for any positive integer y,

$$P[Y > y] = (1-p)^{2y} = \left((1-p)^2\right)^y = \left(1 - (1 - (1-p)^2)\right)^y$$

This shows that Y is geometric with parameter $(1 - (1-p)^2)$. Now, since Y is geometric, its expectation is given by

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

You can solve the problem by directly calculating the mass function of Y also.

$$P[Y = y] = P[X_1 = y, X_2 > y] + P[X_2 = y, X_1 > y] + P[X_1 = y, X_2 = y]$$

$= (1-p)^{y-1}p(1-p)^y + (1-p)^{y-1}p(1-p)^y + (1-p)^{y-1}p(1-p)^{y-1}p$

Hence,

$$f_Y(y) = 2p(1-p)^{2y-1} + p^2(1-p)^{2y-2} = p(1-p)^{2y-2}(2(1-p)+p) = p(1-p)^{2y-2}(2-p)$$

$$= \left((1-p)^2\right)^{y-1} p(2-p) = \left((1-p)^2\right)^{y-1}(1-(1-p)^2)$$

This shows that Y has geometric distribuion with parameter $(1-(1-p)^2)$.

3. (a) A rod of length 1 is broken at a random point. The piece containing the left end is once again broken at a random point. Let L be the length of the final piece containing the left end. Find $E[L]$.

Answer: Let us take the left end of rod as origin. Let X_1 denote the point where the rod is broken the first time. Then X_1 is uniform over $(0, 1)$. Let X_2 denote the point where the rod is broken on the second time. Note that (since we took left end as origin) X_2 is also the length of the final piece containing the left end and hence $L = X_2$. Now, given X_1, we know X_2 is uniform over $(0, X_1)$. What we mean by this is

$$f_{X_2|X_1}(x_2|x_1) = \frac{1}{x_1}, \quad 0 < x_2 < x_1 < 1$$

What we need is $EL = EX_2$. From the above conditional density, it is easy to see that $E[X_2|X_1] = \frac{X_1}{2}$. Hence,

$$E[L] = E[X_2] = E\left[E[X_2|X_1]\right] = E\left[\frac{X_1}{2}\right] = \frac{1}{4}$$

because $EX_1 = \frac{1}{2}$.

(b) Let $X_1, X_2, \cdots$ be *iid* random variables having exponential density with parameter λ. Let N be a geometric random variable with parameter p. N is independent of all X_i. Let $S = X_1 + X_2 + \cdots + X_N$. Find ES and density of S

(You can use the following fact: If X has gamma density with parameters α_1 & λ, and Y has gamma density with parameters α_2 & λ, and X, Y are independent then $X + Y$ is gamma with parameters $\alpha_1 + \alpha_2$ & λ)

Answer: Using the formula for expectation of random number of random variables, we get

$$ES = EN\ EX_1 = \frac{1}{p}\frac{1}{\lambda} = \frac{1}{\lambda p}$$

because N is geometric and X_i are iid exponential.

We can find the density of S as follows. First let us derive an expression for the distribution function of S. For this, first consider

$$P\left[\sum_{i=1}^{N} X_i \le x | N = n\right] = \frac{P[\sum_{i=1}^{N} X_i \le x, N = n]}{P[N = n]}$$

$$= \frac{P[\sum_{i=1}^{n} X_i \le x, N = n]}{P[N = n]} = P\left[\sum_{i=1}^{n} X_i \le x\right]$$

because N and X_i are independent. We know that exponential is a special case of gamma with parameters 1 and λ. So, sum of n iid exponential random variables would be gamma with parameters n and λ. Let $F_{G(n,\lambda)}$ denote this distribution function and let $f_{G(n,\lambda)}$ denote the corresponding density. Then, from the above, we have

$$P\left[\sum_{i=1}^{N} X_i \le x | N = n\right] = P\left[\sum_{i=1}^{n} X_i \le x\right] = F_{G(n,\lambda)}(x)$$

Hence

$$P\left[\sum_{i=1}^{N} X_i \le x\right] = \sum_{n} P\left[\sum_{i=1}^{N} X_i \le x | N = n\right] P[N = n] = \sum_{n=1}^{\infty} f_N(n) F_{G(n,\lambda)}(x)$$

Differentiating this, we get the density of S as

$$f_S(x) = \sum_{n=1}^{\infty} f_N(n) f_{G(n,\lambda)}(x) = \sum_{n=1}^{\infty} (1-p)^{n-1} p \frac{1}{(n-1)!} \lambda^n x^{n-1} e^{-\lambda x}$$

where we have substituted for the gamma density and the mass function of a geometric random variable. We can simplify this as follows.

$$f_S(x) = \sum_{n=1}^{\infty} (1-p)^{n-1}\ p\ \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} = \lambda\ p\ e^{-\lambda x} \sum_{n=1}^{\infty} \frac{((1-p)\lambda x)^{n-1}}{(n-1)!}$$

$$= \lambda \, p \, e^{-\lambda x} \sum_{m=0}^{\infty} \frac{((1-p)\lambda x)^m}{m!} = \lambda \, p \, e^{-\lambda x} \, e^{(1-p)\lambda x} = \lambda \, p \, e^{-\lambda p x}$$

Thus, S is exponential with parameter λp. (Now we can easily see that its expectation is what we got earlier through the formula).

We can also solve this problem using the moment generating functions. The moment generating function of S is

$$M_S(t) = E\left[e^{t(X_1+X_2+\cdots+X_N)}\right]$$

Now we have

$$E\left[e^{t(X_1+X_2+\cdots+X_N)}|N=n\right] = E\left[e^{t(\sum_{i=1}^n X_i)}|N=n\right] = E\left[e^{t(\sum_{i=1}^n X_i)}\right] = \prod_{i=1}^n E\left[e^{tX_i}\right]$$

because X_i are independent of N and X_i are iid. Since X_i are exponential

$$M_{X_i}(t) = E\left[e^{tX_i}\right] = \frac{\lambda}{\lambda - t}, \; t < \lambda$$

Hence we get

$$E\left[e^{t(X_1+X_2+\cdots+X_N)}|N=n\right] = \left(\frac{\lambda}{\lambda - t}\right)^n$$

Thus, we get the moment generating function of S as

$$M_S(t) = E\left[\; E\left[e^{t(X_1+X_2+\cdots+X_N)}|N\right]\right] = E\left[\left(\frac{\lambda}{\lambda-t}\right)^N\right] = \sum_{k=1}^{\infty}(1-p)^{k-1} \, p \, \left(\frac{\lambda}{\lambda-t}\right)^k$$

$$= p\left(\frac{\lambda}{\lambda-t}\right)\sum_{k=1}^{\infty}(1-p)^{k-1}\left(\frac{\lambda}{\lambda-t}\right)^{k-1}$$

What we have here is an infinite geometric series. That would converge if $\frac{\lambda(1-p)}{\lambda-t} < 1$ which is true if $t < \lambda - \lambda(1-p) = \lambda p$. That would be the range of t for which the moment generating function of S exists. Now, summing the infinite series, we get

$$M_S(t) = p\,\frac{\lambda}{\lambda-t}\,\frac{1}{1-\frac{\lambda(1-p)}{\lambda-t}} = \frac{p\lambda}{\lambda-t}\,\frac{\lambda-t}{(\lambda-t)-\lambda(1-p)} = \frac{\lambda p}{\lambda p - t}$$

This is the moment generating function of S and it exists for $t < \lambda p$. Hence S is exponential with parameter λp.

Comment: This is a standard result and it is worth remembering.

4. (a). Let X be Gaussian with mean zero and variance 1. Let Z be a discrete random variable that is independent of X and suppose $\text{Prob}[Z = 1] = \text{Prob}[Z = -1] = 0.5$. Let $Y = ZX$. Find density of Y. Are X, Y uncorrelated? Are X, Y jointly Gaussian?

Answer: We are given $Y = ZX$. The distribution function of Y is given by

$$P[Y \leq y] = P[ZX \leq y] = P[ZX \leq y|Z = 1]P[Z = 1]+P[ZX \leq y|Z = -1]P[Z = -1]$$

$$= \frac{1}{2}\left(P[X \leq y] + P[-X \leq y]\right) = \frac{1}{2}\left(P[X \leq y] + P[X \geq -y]\right)$$

Since X is standard Gaussian, its distribution function is Φ. So, we get

$$F_Y(y) = \frac{1}{2}\left(\Phi(y) + (1 - \Phi(-y))\right) = \Phi(y)$$

(Recall that $\Phi(-x) = 1 - \Phi(x)$). Hence Y is also Gaussian with mean zero and variance 1.

From above, we have $EX = EY = 0$. Now, since X and Z are independent, $EXY = EZX^2 = EZ\ EX^2 = 0$ because $EZ = 0$. Hence, X, Y are uncorrelated.

But X and Y are not independent. (Obvious because Y can be only X or $-X$. If you are not convinced, then: $P[Y > 2|X \in [-1, 1]] = 0$ but $P[Y > 2] \neq 0$!) If X, Y are jointly Gaussian then their uncorrelatedness should imply independence. Hence, X, Y are not jointly Gaussian.

(b). Suppose we have two decks of n cards each. Cards of each deck are numbered $1, 2, \ldots, n$. The two decks are separately shuffled and then the corresponding cards in each deck are compared one by one. We say a match has occurred at position i if the i^{th} card in each deck has the same number. Let S_n denote the total number of matches. Find $E[S_n]$ and $\text{Var}(S_n)$.

Answer: Let $X_1, \cdots X_n$ be indicator random variables with $X_i = 1$ if there is a match at position i. It is easy to see that $P[X_i = 1] = \frac{(n-1)!}{n!} = \frac{1}{n}$. The number of matches is given by $S_n = \sum_{i=1}^{n} X_i$. Hence $ES_n = n\,\frac{1}{n} = 1$.

To calculate the variance, note that X_i are not independent. But we have

$$P[X_i = 1, X_j = 1] = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

Hence

$$\mathrm{Cov}(X_i, X_j) = E[X_i X_j] - EX_i\ EX_j = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}$$

Also,

$$\mathrm{Var}(X_i) = \frac{1}{n}\left(1 - \frac{1}{n}\right)$$

Hence

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) - \sum_{i=1}^{n} \sum_{j \neq i} \mathrm{Cov}(X_i, X_j) = n\frac{1}{n}\left(1 - \frac{1}{n}\right) + n(n-1)\frac{1}{n^2(n-1)} = 1$$

Comment: I hope you can see the similarity with the problem of n men putting their hats together and making random selection. You can think of each man and his hat being given a number and then we randomly pair men with hats. Thus the matches in this problem correspond to man getting his own hat.

E1 222 Stochastic Models and Applications
Test III

`Time: 90 minutes` `Max. Marks:40`
`Date: 17 Nov 2019`

Answer **ALL** questions. All questions carry equal marks
Answers should be written only in the space provided.

1. a. Consider a Markov chain with the following transition probability matrix:

$$P = \begin{bmatrix} 0.25 & 0.2 & 0 & 0.3 & 0.25 \\ 0 & 0.25 & 0.75 & 0 & 0 \\ 0 & 0.75 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0.7 & 0.3 \end{bmatrix}$$

Specify which are the transient and recurrent states and find all the closed irreducible subsets of recurrent states. Find the absorption probabilities from each of the transient states to each of the closed irreducible subsets of recurrent states.

Answer: Let us take the state space as $\{0, 1, 2, 3, 4\}$. It is easy to see that $\{1, 2\}$ is a closed irreducible set. (The two states communicate with each other and from 1 or 2 we cannot go to any other states). Similarly $\{3, 4\}$ is a closed irreducible set. Since from state 0 we can go to state 1 or state 3, it is a transient state. Hence, the decomposition of state space is as follows: $S_T = \{0\}$, $S_R = \{1, 2\} + \{3, 4\}$.
Let $C = \{1, 2\}$. Since there is only one transient state we need to compute $\rho_C(0)$. We have

$$\rho_C(0) = P(0, 1) + P(0, 2) + P(0, 0)\rho_C(0) = 0.2 + 0.25\rho_C(0)$$

This gives us $\rho_C(0) = 0.2/0.75 = 4/15$.
Let $C' = \{3, 4\}$. Then $\rho_{C'}(0) = 1 - \rho_C(0) = 11/15$.

b. Define positive recurrent and null recurrent states in a Markov chain. If π is a stationary distribution of the chain and i is a null recurrent state then explain why we must have $\pi(i) = 0$. Explain why a finite Markov chain cannot have a null recurrent state.

Answer: Let y be a recurrent state. Let T_y be the hitting time (or first passage time). That is, $T_y = \min\{n : n > 0, X_n = y\}$. Let $m_y = E_y[T_y]$ where E_y is expectation conditioned on $X_0 = y$.

The recurrent state y is called positive recurrent if $m_y < \infty$; otherwise (that is, if $m_y = \infty$) it is called null recurrent.

Let $N_n(y)$ denote the number of visitis to y till time n. Let $G_n(x, y) = E_x[N_n(y)]$. Then we have the result that $\frac{G_n(x,y)}{n} \to \frac{\rho_{xy}}{m_y}$ as $n \to \infty$. Hence, if y is null recurrent, then, $\frac{G_n(x,y)}{n} \to 0$.

If we let $I_y(X_n)$ be the indicator for X_n being y, then, $N_n(y) = \sum_{m=1}^{n} I_y(X_m)$. Since $E_x[I_y(X_n)] = P^n(x, y)$, we have $G_n(x, y) = \sum_{m=1}^{n} P^m(x, y)$. Thus

$$\frac{G_n(x, y)}{n} = \frac{1}{n} \sum_{m=1}^{n} P^m(x, y)$$

Hence, if y is null recurrent then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x, y) = 0$$

If π is a stationary distribution, then we have $\pi(y) = \sum_x \pi(x)P^m(x, y)$ for all m. Summing this over $m = 1, \cdots, n$ and dividing by n, we get

$$\pi(y) = \frac{1}{n} \sum_{m=1}^{n} \sum_{x} \pi(x)P^m(x, y) = \sum_{x} \pi(x) \frac{1}{n} \sum_{m=1}^{n} P^m(x, y), \forall n$$

Since this holds for all n, it also holds in the limit. The limit of the RHS above, as $n \to \infty$, is zero if y is null recurrent, as explained earlier. (Note that we can take the limit inside the summation over x because of bounded convergence theorem). Thus, if y is null recurrent and π is a stationary distribution, then $\pi(y) = 0$.

Suppose, A is a finite closed set. Since it is closed, $\sum_{j\in A} P^m(i, j) = 1$, $\forall i \in A$, $\forall m$. Hence, we have

$$1 = \frac{1}{n} \sum_{m=1}^{n} \sum_{j\in A} P^m(i, j) = \sum_{j\in A} \frac{1}{n} \sum_{m=1}^{n} P^m(i, j), \forall n$$

Since the above holds for all n, it also holds as $n \to \infty$. Suppose all states in A are null recurrent. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(i,j) = 0, \forall j \in A$$

Thus, if all states in the finite closed set A are null recurrent, then,

$$1 = \lim_{n\to\infty} \sum_{j\in A} \frac{1}{n} \sum_{m=1}^{n} P^m(i,j) = \sum_{j\in A} \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(i,j) = 0$$

which is a contradiction. In the above, we could take the limit inside the outer summation because A is a finite set.

The above shows that in a finite closed set, not all states can be null recurrent.

We have a result that if x leads to y and x is positive recurrent then y is positive recurrent. Hence, in a closed irreducible set of recurrent states, either all states positive recurrent or all states are null recurrent.

Suppose a finite chain has a null recurrent state. That state must be in one of the closed irreducible sets of recurrent states. In a finite chain, all closed irreducible sets of recurrent states have to be finite and hence none of them can be wholly null recurrent. Thus, a finite chain cannot have a null recurrent state.

2. a. Let $\{X_n,\ n \geq 0\}$ be an irreducible, positive recurrent aperiodic Markov chain whose stationary probabilities are given by π. Define another process $\{Y_n,\ n \geq 1\}$ by $Y_n = (X_{n-1}, X_n)$. (That is, Y_n keeps track of the last two states of the original chain). Is Y_n a Markov Chain? If so, find its transition probabilities and $\lim_{n\to\infty} \text{Prob}[Y_n = (i,j)]$.

Answer: Given $\{X_n\}$ is a Markov chain. Hence, conditioned on X_{n-2}, X_n, X_{n-1} are conditinally independent of $X_{n-3}, X_{n-4}, \cdots$. Hence, it is easy to see that $\{Y_n\}$ is a Markov chain.

We can calculate the transition probabilities as follows:

$$\begin{aligned}
\text{Prob}[Y_n = (i,j) | Y_{n-1} = (k,l)] &= \text{Prob}[X_{n-1} = i, X_n = j | X_{n-2} = k, X_{n-1} = l] \\
&= \begin{cases} 0 & \text{if } i \neq l \\ \text{Prob}[X_n = j | X_{n-1} = i] & \text{otherwise} \end{cases}
\end{aligned}$$

We can calculate the limiting probabilities as follows. Since $\{X_n\}$ is an irreducible positive recurrent aperiodic chain withstationary probabilities π, we have $\lim_{n\to\infty} \text{Prob}[X_n = i] = \pi(i)$

$$\begin{aligned}
\lim_{n\to\infty} \text{Prob}[Y_n = (i,j)] &= \lim_{n\to\infty} \text{Prob}[X_{n-1} = i, X_n = j] \\
&= \lim_{n\to\infty} \text{Prob}[X_n = j | X_{n-1} = i]\text{Prob}[X_{n-1} = i] \\
&= P(i,j)\pi(i)
\end{aligned}$$

where $P(i,j)$ are the transition probabilities of X_n.

b. Define transient and recurrent states in a Markov chain. When is a Markov chain said to be irreducible. Can an irreducible markov chain with countably infinite state space have some transient states and some recurrent states?

Answer: For any state y the hitting time is $T_y = \min\{n : n > 0, X_n = y\}$. Define $\rho_{xy} = \text{Prob}[T_y < \infty | X_0 = x]$. ρ_{xy} is the probability that starting in x, sometime or the other we hit y.

A state y is defined to be transient if $\rho_{yy} < 1$ and is defined to be recurrent if $\rho_{yy} = 1$.

A Markov chain is said to be irreducible if x leads to y for all x, y. That is, $\rho xy > 0$ for all states x, y.

We have a result that if x is recurrent and x leads to y then y is recurrent. In an irreducible chain, every state leads to every other state. Hence, in an irreducible chain, if any one state is recurrent then all states are recurrent. Thus, in an irreducible chain either all states are transient or all states are recurrent; we cannot have some transient and some recurrent states.

3. a. Let $X_1, X_2, \cdots$ be a sequence of discrete random variables with X_n being geometric with parameter λ/n where we have $0 < \lambda < 1$. Let $Z_n = X_n/n$. Does Z_n converges in distribution?

Answer: We have

$$F_{Z_n}(x) = P[Z_n \leq x] = P[X_n \leq nx] = 1 - \left(1 - \frac{\lambda}{n}\right)^{[nx]}$$

where $[nx]$ is the greatest integer smaller than or equal to nx. We need to find the limit of the RHS above as $n \to \infty$. For that we have

$$\left(1-\frac{\lambda}{n}\right)^{[nx]} = \left(1-\frac{\lambda}{n}\right)^{nx} \left(1-\frac{\lambda}{n}\right)^{[nx]-nx}$$

For all n, $[nx]-nx$ is between -1 and zero and hence, as $n \to \infty$, the second factor on the RHS above goes to 1. The first term goes to $e^{-\lambda x}$. Thus,

$$\lim_{n\to\infty} F_{Z_n}(x) = 1 - e^{-\lambda x}$$

which shows that Z_n converges in distribution to the exponential distribution.

b. Let $X_1, X_2, \cdots$ be a sequence of independent random variables with

$$\text{Prob}[X_n = n] = \text{Prob}[X_n = -n] = \frac{1}{2n^2}; \quad \text{Prob}[X_n = 0] = \frac{n^2-1}{n^2}$$

Does this sequence converge (i) in probability, (ii) with probability one, (iii) in the r^{th} mean?

Answer: The probability of X_n being zero is increasing with n and hence a good candidate for limit is zero. To test convergence in probability, we have

$$\lim_{n\to\infty} P[|X_n - 0| > \epsilon] = \lim_{n\to\infty} P[X_n = n \text{ or } X_n = -n] = \lim_{n\to\infty} \frac{2}{2n^2} = 0$$

Hence $X_n \xrightarrow{P} 0$.

Now, if X_n converges in any other mode, it should converge to zero. To test for convergence almost surely, we can use Borel-Cantelli lemma. Let $A_n^\epsilon = [|X_n| > \epsilon]$. Then, as calculated above, $P[A_n^\epsilon] = \frac{1}{n^2}$, $\forall \epsilon > 0$. Thus, for all $\epsilon > 0$, $\sum_n P[A_n^\epsilon] < \infty$. Hence, by Borel-Cantelli lemma, X_n converges to zero almost surely.

To test for convergence in r^{th} mean we have

$$E[|X_n|^r] = (|n|^r + |-n|^r)\,\frac{1}{2n^2} = n^{r-2}$$

Hence, X_n converges in r^{th} mean for $r < 2$ and it does not converge for $r \geq 2$.

4. a. Twenty real numbers are rounded off to the nearest integer and added. Assume that the individual rounding-off errors are independent and uniformly distributed over $(-0.5,\ 0.5)$. Find the (approximate) probability that the sum obtained like this differes from the sum of the original twenty numbers by more than 3.

Answer: Let $X_1, \cdots, X_{20}$ represent the roundoff errors in the twenty numbers. It is given that X_i are iid uniform over $(-0.5,\ 0.5)$. Hence, $EX_i = 0$ and $\mathrm{Var}(X_i) = 1/12$. Let $S = \sum_{i=1}^{20} X_i$. Then, $ES = 0$ and $\mathrm{Var}(S) = 20/12 = 5/3$. The difference between the original sum and the sum of rounded-off numbers is $|S|$. Hence, the probability we want is

$$\begin{aligned} P[|S| > 3] &= P[S < -3] + P[S > 3] \\ &= P\left[\frac{S}{\sqrt{5/3}} < \frac{-3}{\sqrt{5/3}}\right] + P\left[\frac{S}{\sqrt{5/3}} > \frac{3}{\sqrt{5/3}}\right] \\ &\approx \Phi\left(\frac{-3}{\sqrt{5/3}}\right) + 1 - \Phi\left(\frac{3}{\sqrt{5/3}}\right) \end{aligned}$$

b. Suppose $X_1, X_2, \cdots$ be *iid* continuous random variables with density $f(x) = 2x,\ 0 \le x \le 1$. Let $S_n = \sum_{i=1}^{n} X_i^2$. Does S_n/n converge almost surely? (Answer Yes/No with a short justification). If your answer is yes, find x_0 such that the sequence S_n/n converges to x_0 with probability one. Also, explain how we can approximately find the probability of $[-a \le \frac{S_n}{n} - x_0 \le b]$ for some $a, b > 0$

Answer: Since X_i are iid, X_i^2 would also be iid. Hence, by strong law of large numbers, S_n/n converges almost surely to EX_1^2.
From the given density of X_i, we get

$$EX_1^2 = \int_0^1 x^2\ 2x\ dx = 0.5$$

Hence $\frac{S_n}{n} \stackrel{a.s.}{\to} 0.5$.
Let σ^2 be the variance of X_i^2. We can calculate it using the density for X_i. (Note that $\sigma^2 = EX_i^4 - (EX_i^2)^2$).
We have: $ES_n = 0.5n$ and $\mathrm{Var}(S_n) = n\sigma^2$.

Now we have

$$
\begin{aligned}
P\left[-a \leq \frac{S_n}{n} - 0.5 \leq b\right] &= P[-na \leq S_n - 0.5n < nb] \\
&= P\left[\frac{-na}{\sigma\sqrt{n}} \leq \frac{S_n - 0.5n}{\sigma\sqrt{n}} \leq \frac{nb}{\sigma\sqrt{n}}\right] \\
&\approx \Phi\left(\frac{nb}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{-na}{\sigma\sqrt{n}}\right)
\end{aligned}
$$

This is how we can approximately find the required probability.

E1 222 Stochastic Models and Applications
Test I

Time: 75 minutes Max. Marks:40
Date: 20 Sept 2021

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

1. a. Two players A and B are playing a game that consists of a series of points. Each point is independently won by A with probability p and by B with probability $1-p$. When one player has won two points more than the other player, the game ends and the winner is the player with more points. Calculate the probability that the game would end after m points are played. Also, calculate the probability that player A wins the game.

Answer: I hope you can easily see that the game ends only after even number of points are played. To see this, suppose when the game ends the two players have K_1 and K_2 points with $K_1 - K_2 = 2$. Then, $K_1 + K_2 = 2K_2 + 2$. So, the probability is non-zero only when $m = 2n$.

If the game ends at $2n$ then, after $2n-2$ points are played they should have equal number of points. (If one player had two points more then the game should have ended at $2n-2$; a player cannot lead by one point after an even number of points are played). Thus, for the game to end after $2n$ points, at the end of $2r$ points, for each of $r = 1, 2, \cdots, n-1$, they should be equal on points and finally the same player should have won the last two points. So, players A, B should win one point each (in any order) out of points 1 and 2, and similarly for points 3 and 4 and so on. Since the points are independent,

$$P[\text{game ends after } 2n \text{ points}] = (2p(1-p))^{n-1}\left(p^2 + (1-p)^2\right)$$

You can verify that

$$\sum_{n=1}^{\infty} (2p(1-p))^{n-1}\left(p^2 + (1-p)^2\right) = \frac{1}{1-2p(1-p)}\ \left((p + (1-p))^2 - 2p(1-p)\right) = 1$$

Now, it is easy to see that

$$P[\text{Player } A \text{ wins the game}] = \sum_{n=1}^{\infty} (2p(1-p))^{n-1}\, p^2 = \frac{p^2}{1-2p(1-p)} = \frac{p^2}{p^2+(1-p)^2}$$

I hope you can see that the final expression is what you expect because the game can be viewed as repeating the random experiment of playing two points till one of the players wins both points.

b. Take $\Omega = \{(x,\ y)\ :\ 0 \leq x, y \leq 1\}$ with the usual probability assignment where the probability of an event is proportional to its area. Every time this random experiment is performed, the outcome is a point in $\Omega \subset \Re^2$. Suppose this random experiment is repeated till the point obtained is such that its distance from the origin is less than 0.5. Let X denote the number of repetitions needed. Find the expected value of X.

Answer: As is easy to see here we are (independently) repeating a random experiment till an event occurs and X is the number of repetitions needed. Hence X would be geometric with parameter equal to the probability of the event. The event is that the point (x, y) should be inside a circle with center at origin and radius 0.5 (and (x, y) is also inside Ω). The area of this subset of Ω is $0.25(\pi(0.5)^2) = \pi/16$. Hence $E[X] = 16/\pi$.

2. a. Let X be a continuous random variable with pdf

$$f_X(x) = \frac{K}{x^4},\ x \geq 3$$

Find value of K, $E[X]$, $\text{Var}(X)$ and $P[X \leq 9]$.

Answer: We have

$$\int_3^{\infty} \frac{K}{x^4}\, dx = K \left. \frac{-1}{3x^3} \right|_3^{\infty} = K \frac{1}{81}$$

Hence, $K = 81$. We can calculate $P[X \leq 9]$ as

$$\int_3^9 \frac{81}{x^4}\, dx = 81 \left. \frac{-1}{3x^3} \right|_3^9 = 81 \left(\frac{1}{81} - \frac{1}{3 * 9^3} \right) = \frac{26}{27}$$

The EX is given by

$$E[X] = \int_3^{\infty} x \frac{81}{x^4}\, dx = 81 \left. \frac{-1}{2x^2} \right|_3^{\infty} = 81 \frac{1}{18} = 4.5$$

We can get $E[X^2]$ as

$$E[X^2] = \int_3^\infty x^2 \frac{81}{x^4}\, dx = 81 \left. \frac{-1}{x} \right|_3^\infty = 81 \frac{1}{3} = 27$$

Hence $\mathrm{Var}(X) = 27 - (4.5)^2 = 6.75$

b. 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,

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

Answer: Since $\Phi_X(1) = 1$, we only need to consider $0 < t < 1$,

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

This gives us the required relation:

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

Comment: Using Φ_X we can also bound $P[X \ge y]$. 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 $t \ge 1$

$$\begin{aligned}
\Phi_X(t) &= E\left[t^X\right] = \sum_{k=0}^{\infty} t^k\, f_X(k) \\
&= \sum_{k=0}^{y-1} t^k\, f_X(k) \ + \sum_{k=y}^{\infty} t^k\, f_X(k)
\end{aligned}$$

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

This gives us the relation

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

Note that these bounds are valid for all t in an interval. Hence, for a specific X we can get a tight bound by choosing a t in that interval to minimize $\frac{\Phi_X(t)}{t^y}$

3. a. A total of m keys are to be put in n boxes with each key independently being put in box-i with probability p_i, $i = 1, \cdots, n$. Every time a key is put in a non empty box, we say a collison has occurred in that box. Find the expected number of collisons in box-1

Answer: Let X denote the number of keys put in box-1. From the description given, we have the following: each key is independently put in box-1 with probability p_1 and in one of the other boxes with probability $1 - p_1$. Then X is binomial with parameters m and p_1. Let Y denote the number of colisions in box-1. A collision occurs everytime we put a key in a non-empty box. Hence we have $Y = X - 1$ if $X > 1$ and $Y = 0$ otherwise. Hence, we have $Y = \max(0, X - 1)$. Since we know X is binomial(m, p_1), we can now compute EY:

$$\begin{aligned}
E[Y] &= \sum_{k=1}^{m-1} kP[Y=k] \\
&= \sum_{k=1}^{m-1} kP[X=k+1] \\
&= \sum_{k=1}^{m-1} (k+1)P[X=k+1] - \sum_{k=1}^{m-1} P[X=k+1] \\
&= \sum_{k=2}^{m} kP[X=k] - \sum_{k=2}^{m} P[X=k]
\end{aligned}$$

$$
\begin{aligned}
&= EX - P[X=1] - (1 - P[X=0] - P[X=1]) \\
&= EX - 1 + P[X=0] \\
&= mp_1 - 1 + (1-p_1)^n
\end{aligned}
$$

Comment: As you can see the only issue in the problem is realizing that the number of collisions is $Y = \max(0, X-1)$ with X binomial. Given this equation, you already know how to calculate EY from your assignment-1.

b. Suppose X is a continuous random variable with pdf

$$f_X(x) = \frac{\lambda e^{-\lambda x}}{(1+e^{-\lambda x})^2}, \quad -\infty < x < \infty$$

where $\lambda > 0$. Define $Y = \frac{1}{1+e^{-\lambda X}}$. Find the pdf of Y.

Answer: Let $g(x) = 1/(1+e^{-\lambda x})$. Note that $g(x) \in (0,1)$. By differentiating g we can easily see that it is monotone. Hence we can use the theorem. First, let us find its inverse. (Note that the inverse is defined for $y \in (0,1)$).

$$y = \frac{1}{1+e^{-\lambda x}} \;\Rightarrow\; e^{-\lambda x} = \frac{1}{y} - 1 \;\Rightarrow\; x = \frac{-1}{\lambda}\ln\left(\frac{1-y}{y}\right)$$

Hence we get

$$\frac{d}{dy}g^{-1}(y) = \frac{-1}{\lambda}\frac{y}{1-y}\frac{-1}{y^2} = \frac{1}{\lambda y(1-y)}$$

Now, using the theorem, we can get density of Y. For that we need to calculate $f_X(g^{-1}(y))$. For that, first note that

$$e^{-\lambda g^{-1}(y)} = e^{-\lambda \frac{-1}{\lambda}\ln\left(\frac{1-y}{y}\right)} = \frac{1-y}{y}$$

Hence

$$f_X(g^{-1}(y)) = \lambda\frac{\frac{1-y}{y}}{(1+\frac{1-y}{y})^2} = \lambda y(1-y)$$

This gives us

$$f_Y(y) = \left|\frac{d}{dy}g^{-1}(y)\right| f_X(g^{-1}(y)) = \frac{1}{\lambda y(1-y)}\lambda y(1-y) = 1, 0 < y < 1$$

This shows that Y is uniform over $(0,1)$.

Comment: You need not do all the calculation shown above for this problem. If you differentiated $g(x)$ to see whether it is monotone, you would have immediately seen that the derivative is the density function of X. (You could have noticed this even just by looking at the two functions). This means that what we have is $Y = g(X)$ where $g(\cdot)$ is the distribution function of X. Hence we know, as shown in class, Y is uniform over $(0, 1)$.

You could have also solved this from first principles: for $0 < y < 1$

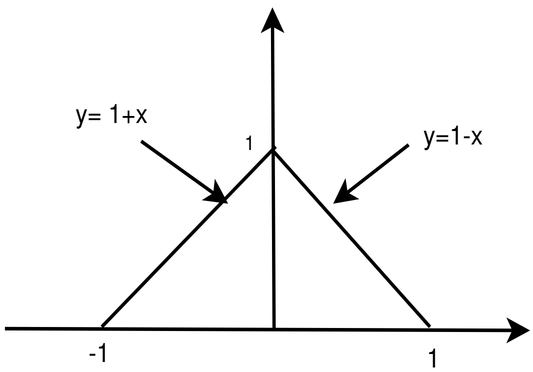
$$\begin{aligned} F_Y(y) &= P\left[\frac{1}{1+e^{-\lambda X}} \leq y\right] \\ &= P[e^{-\lambda X} \geq \frac{1}{y} - 1] \\ &= P\left[-\lambda X \geq \ln\left(\frac{1-y}{y}\right)\right] \\ &= P\left[X \leq \frac{-1}{\lambda} \ln\left(\frac{1-y}{y}\right)\right] \end{aligned}$$

Now you can complete the problem by integrating the density function of X.

Comment: This is a problem where some of you made bad errors of notation. You cannot write the formula as $f_Y(y) = J f_X(x)$. We assume J represents the needed derivative. But what is the 'symbol' x on the RHS? Suppose I define a function $h : \Re \to \Re$ as $h(x) = 3 + z$. This is a meaningless definition of a function unless I specify what z is in terms of x. That is the reason we have to always write the argument of f_X in the formula as $g^{-1}(y)$.

4. a. Let X, Y be continous random variables with joint density, $f_{XY}(x, y) = 1$ if (x, y) is inside the triangle with vertices at $(-1, 0)$, $(0, 1)$ and $(1, 0)$. (Otherwise f_{XY} is zero). Find the marginal densities f_X and f_Y and $P[Y > 0.6 \mid X = -0.1]$.

Answer: A figure will help you understand the problem better.



The region is $\{(x,y) : -1 \le x \le 1, 0 \le y \le 1-|x|\}$ or $\{(x,y) : 0 \le y \le 1, y-1 \le x \le 1-y\}$. The density $f_{XY}(x,y)$ is 1 if (x,y) is in this region and is zero otherwise.

Now, we can get the marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy = \int_0^{1-|x|} dy = 1-|x|, \; -1 \le x \le 1$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx = \int_{y-1}^{1-y} dx = 2(1-y), \; 0 \le y \le 1$$

To calculate the required conditional probability, we need $f_{Y|X}$.

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{1}{1-|x|}, \; -1 \le x \le 1, 0 \le y \le 1-|x|$$

Now we get

$$P[Y > 0.6|X = -0.1] = \int_{0.6}^{\infty} f_{Y|X}(y|-0.1)\, dy = \int_{0.6}^{0.9} \frac{1}{0.9}\, dy = \frac{0.3}{0.9} = \frac{1}{3}$$

From the expression for $f_{Y|X}$ we can see that, conditioned on X, the Y is uniform over $(0, 1-|X|)$. Thus, given that $X = -0.1$ we know that Y is uniform over $(0, 0.9)$ and hence we know the probability that $Y > 0.6$ is $1/3$.

b. Let X, Y be iid random variables having geometric distribution with parameter p. (That is, X, Y are independent and both have the mass function $f(i) = (1-p)^{i-1}p, \; i = 1, 2, \cdots$). Let $Z = X+Y$. Find the conditional mass function $f_{X|Z}$. (Hint: You need not explicitly calculate the mass function of Z).

Answer: Given that X, Y are iid geometric(p). Now we have

$$\begin{aligned} f_{X|Z}(x|z) &= P[X = x \mid Z = z] \\ &= \frac{P[X = x, Z = z]}{P[Z = z]} \\ &= \frac{P[X = x, Z = z]}{\sum_x P[X = x, Z = z]} \\ &= \frac{P[X = x, X + Y = z]}{\sum_x P[X = x, X + Y = z]} \\ &= \frac{P[X = x, Y = z - x]}{\sum_x P[X = x, Y = z - x]} \\ &= \frac{(1-p)^{x-1}p\ (1-p)^{z-x-1}p}{\sum_{x=1}^{z-1}(1-p)^{x-1}p\ (1-p)^{z-x-1}p},\ z = 2, 3, \cdots,\ x = 1, 2, z - 1 \end{aligned}$$

because $X, Y \geq 1$ and $Z = X + Y$

$$\begin{aligned} &= \frac{(1-p)^{z-2}p^2}{\sum_{x=1}^{z-1}(1-p)^{z-2}p^2} \\ &= \frac{1}{z-1} \end{aligned}$$

Thus, X, conditined on Z is uniform over $\{1, 2, \cdots, Z - 1\}$.

As you can see, from the above algebra we can also see that

$$P[Z = z] = \sum_{x=1}^{z-1}(1-p)^{z-2}p^2 = p^2(1-p)^{z-2}(z-1),\ z = 2, 3, \cdots$$

(Can you show that the above is a pmf?)

We need not explicitly calculate this first, to get our conditional mass function.

E1 222 Stochastic Models and Applications
Test II

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

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

1. a. Let X, Y be continuous random variables with joint density

$$f_{XY}(x,y) = \frac{1}{1-x}, \quad 0 < x < y < 1$$

Find $E[X], E[Y]$ and $E[Y|X]$

Answer: We can get the marginal densities as follows

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\; dy = \int_x^1 \frac{1}{1-x}\; dy = \frac{1}{1-x} \quad y|_x^1 = 1, \; 0 < x < 1 \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\; dx = \int_0^y \frac{1}{1-x}\; dx = -\int_1^{1-y} \frac{1}{t}\; dt = -\log(1-y), \; 0 < y < 1 \end{aligned}$$

Since X is uniform over $(0,\ 1)$, we know $EX = \frac{1}{2}$. We can get the conditional density $f_{Y|X}$ as

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{1}{1-x}, \quad 0 < x < y < 1$$

Conditioned on X, Y is uniform over $(X,\ 1)$. Hence we get $E[Y|X] = \frac{1+X}{2}$. We can also get it directly as follows

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\; dy = \int_x^1 y\; \frac{1}{1-x}\; dy = \frac{1}{1-x}\left(\frac{1}{2} - \frac{x^2}{2}\right) = \frac{1+x}{2}$$

Since we know EX and $E[Y|X]$, we can get EY as

$$E[Y] = E[\ E[Y|X]\ ] = E\left[\frac{1+X}{2}\right] = \frac{1+EX}{2} = \frac{3}{4}$$

As you can see, we do not need to even calculate density of Y. We can also get EY directly from the density of Y

$$\begin{aligned} E[Y] &= -\int_0^1 y\; \log(1-y)\; dy = \int_1^0 (1-t)\; \log(t)\; dt = -\int_0^1 \log(t) dt + \int_0^1 t\log(t)\; dt \\ &= 1 + \left(\frac{t^2}{2}\log(t)\Big|_0^1 - \int_0^1 \frac{t}{2}\; dt\right) = 1 - \frac{1}{4} = \frac{3}{4} \end{aligned}$$

b. Let X, Y be discrete random variables, taking non-negative integer values, with joint mass function

$$P[X=i, Y=j] = e^{-(a+bi)} \frac{(bi)^j \; a^i}{j! \; i!}$$

Find $\mathrm{Cov}(X, Y)$.

Answer: We can rewrite the joint mass function as

$$P[X=i, Y=j] = e^{-a} \frac{a^i}{i!} \; e^{-bi} \frac{(bi)^j}{j!}$$

Hence, X is Poisson with parameter a and Y, conditioned on X, is Poisson with parameter bX.

We can explicitly calculate them also as follows

$$P[X=i] = \sum_{j=0}^{\infty} e^{-a} \frac{a^i}{i!} \; e^{-bi} \frac{(bi)^j}{j!} = e^{-a} \frac{a^i}{i!} \sum_{j=0}^{\infty} e^{-bi} \frac{(bi)^j}{j!} = e^{-a} \frac{a^i}{i!}$$

and hence

$$f_{Y|X}(j|i) = P[Y=j|X=i] = \frac{P[X=i, Y=j]}{P[X=i]} = e^{-bi} \frac{(bi)^j}{j!}$$

I hope all of you are able to see this without doing the above calculation.

Hence, we have $EX = a$ and $E[Y|X] = bX$. Thus, $EY = ab$.

$$E[XY] = E[\; E[XY|X] \;] = E[X \; E[Y|X] \;] = E[bX^2] = b(a+a^2) = ab+a^2b$$

You can calculate $E[XY] = \sum_{i,j} i \; j \;\; P[X=i, Y=j]$ but that would need more involved calculation. Now we get

$$\mathrm{Cov}(X, Y) = E[XY] - E[X] \; E[Y] = ab + a^2b - a^2b = ab$$

Marks for part(a) – 5
Marks for part(b) – 5

2. a. Let X, Y be iid continuous random variables having exponential distribution with $\lambda = 1$. Let $Z = X + Y$ and $W = \frac{X}{X+Y}$. Show that Z and W are independent.

Answer: The given transformation is invertible and the inverse transform is $X = ZW,\ Y = Z - ZW$. The jacobian of the inverse transform is

$$\begin{vmatrix} w & z \\ 1-w & -z \end{vmatrix} = -wz - z + zw = -z$$

Hence the joint density of Z, W is

$$f_{ZW}(z, w) = |z|\ f_{XY}(zw, z - zw)$$

Since both X, Y are exponential, for the joint density above to be non-zero we need $zw > 0$ and $z > zw$. This gives us $z > 0$ and $0 < w < 1$. Hence, the joint density is

$$f_{ZW}(z, w) = z\ e^{-zw}\ e^{-(z-zw)} = z\ e^{-z}, \quad z > 0,\ 0 < w < 1$$

From this joint density, we can immediately see the marginals as: Z is Gamma and W is uniform over $(0,\ 1)$ and the joint is product of the marginals. We can get this also by explicit calculation as

$$f_Z(z) = \int_0^1 z\ e^{-z}\ dw = z\ e^{-z},\ z > 0; \quad \text{and} \quad f_W(w) = \int_0^\infty z\ e^{-z} = 1,\ 0 < w < 1$$

This shows that Z, W are independent because the joint density is the product of the marginals.

Comment: This shows that $\frac{X}{X+Y}$ is uniform over $(0,\ 1)$ when X, Y are iid exponential, which is a useful general result.

b. Let $X_1, X_2, \cdots$ be *iid* continuous random variables. We say a record has occurred at m if $X_m > \max(X_{m-1}, \cdots, X_1)$. Let $N = \min\{n\ :\ n > 1 \text{ and a record occurs at time } n\}$. Show that $EN = \infty$.

Answer: We essentially need the mass function of N to get the expectation. Given the definition of N, it is easier to calculate $P[N > m]$. The event $[N > 2]$ implies that there is no record at time 2 which implies $X_2 < X_1$. Similarly $[N > 3]$ implies $X_3 < \max(X_1, X_2) = X_1$ because there is no record at 3 or 2. Thus, it is easy to see that the event of there being no record up to and including m, is same as $X_1 > X_i,\ i \le m$. Since all possible orderings of iid continuous random variables are equally likely (see answer to problem 4 in Assignment-3), we get

$$P[N > m] = \frac{(m-1)!}{m!} = \frac{1}{m}$$

Since N is a positive integer valued random variable, we get

$$E[N] = \sum_{k=1}^{\infty} P[N > k] = \sum_{k=1}^{\infty} \frac{1}{k} = \infty$$

Comment: You have proved the above formula in one of your assignments. Since you have $P[N > k]$ we can also calculate its mass function as $P[N = k] = \frac{1}{k-1} - \frac{1}{k} = \frac{1}{k(k-1)}$. Now, $EN = \sum_{k=2}^{\infty} k \; \frac{1}{k(k-1)} = \sum_{k=2}^{\infty} \frac{1}{(k-1)} = \infty$

You can also use the idea of the indicator random variables that we mentioned for Q8 in problem sheet 3.6. The event of $[N = k]$ is that there is no record at2, no record at 3, and so on till $k-1$ and there is a record at k. Using those indicator random variables and their independence this probability is $(1-1/2)(1-1/3)\cdots(1-1/(k-1))(1/k) = \frac{1}{k(k-1)}$ which is same as above.

Marks for part(a) – 5

Marks for part(b) – 5

3. a. Consider repeated independent tosses of a coin whose probability of heads is p, $0 < p < 1$. Let X denote the number of tosses needed to get at least one head and one tail. Let Y denote the number of tosses needed to get a head immediately followed by a tail. Find EX and EY.

Answer: Let Z be the indicator random variable of whether or not the first toss was head. For EX, we can condition on Z. If first toss is a head then we need to wait for a tail and the expected number of additional tosses for it would be $\frac{1}{1-p}$ and if the first toss is a tail we need to wait for a head. Thus

$$\begin{aligned} EX &= E[\; E[X|Z] \;] \\ &= E[X|Z=1]\; p + E[X|Z=0]\; (1-p) \end{aligned}$$

$$
\begin{aligned}
&= \left(1+\frac{1}{1-p}\right)p+\left(1+\frac{1}{p}\right)(1-p) \\
&= 1+\frac{p}{1-p}+\frac{1-p}{p} \\
&= \frac{p(1-p)+p^2+(1-p)^2}{p(1-p)} = \frac{(p+1-p)^2-p(1-p)}{p(1-p)} = \frac{1}{p(1-p)}-1
\end{aligned}
$$

To calculate EY. We have to first wait for a head. Then we have to wait for a tail. (After we get a head, till we get a tail everything would be heads and hence the first tail would be the first time we get a head followed by a tail). Hence we get

$$E[Y] = \frac{1}{p}+\frac{1}{1-p} = \frac{1}{p(1-p)}$$

Comment: You had solved a problem in one of the problem sheets about expected number of rolls of a fair dice to get all numbers at least once. There we said that you first wait for any of the 6 numbers then any of the five numbers and so on and thus the expected number of rolls is $\frac{6}{6}+\frac{6}{5}+\frac{6}{4}+\cdots$. Can we use a similar argument here for EX? Initially we wait for any of the two outcomes and then for the remaining one. But we do not know the expected number of tosses for the 'remaining' one because we do not know if the remaining one is head or tail. That is why we needed to condition on the first toss. If the coin is fair, then we can get the answer as $\frac{1}{1}+\frac{2}{1}=3$ which is same as what the above formula gives for $p=0.5$

In this problem, it is not difficult to find the distributions of the random variables. For $[X=k]$, by k^{th} toss we have got one head and one tail for the first time. So, either k^{th} toss is a head and everything before it is a tail or it is a tail and everything before it is a head. Hence

$$P[X=k] = (1-p)^{k-1}p + p^{k-1}(1-p),\ k=2,3,\cdots$$

Now we can calculate EX as

$$EX \;=\; \sum_{k=2}^{\infty} k\,(1-p)^{k-1}\,p \;+\; \sum_{k=2}^{\infty} k\,p^{k-1}\,(1-p)$$

$$
\begin{aligned}
&= (1-p)\sum_{k=2}^{\infty} k\,(1-p)^{k-2}\,p \;+\; p\sum_{k=2}^{\infty} k\,p^{k-2}\,(1-p) \\
&= (1-p)\sum_{k'=1}^{\infty} (k'+1)\,(1-p)^{k'-1}\,p \;+\; p\sum_{k'=1}^{\infty} (k'+1)\,p^{k'-1}\,(1-p) \\
&= (1-p)\,\left(\frac{1}{p}+1\right) \;+\; p\left(\frac{1}{1-p}+1\right) = 1 + \frac{1-p}{p} + \frac{p}{1-p}
\end{aligned}
$$

Similarly we can find mass function of Y also. What is the event of $Y = k$? That should mean we got a Tail on k^{th} toss and a head on toss number $k-1$. What about tosses 1 to $k-2$? If there is any tail in these then all the tails have to be in the beginning because we cannot have a tail after a head in these tosses. Thus we can have s tails in the beining followed by heads till toss $k-2$ and s can be between 0 and $k-2$. Hence we have, for $k = 2, 3, \cdots$,

$$
\begin{aligned}
P[Y = k] &= \sum_{s=0}^{k-2} (1-p)^s p^{k-2-s} p(1-p) \\
&= p^{k-1}\,(1-p)\,\sum_{s=0}^{k-2}\left(\frac{1-p}{p}\right)^s \\
&= p^{k-1}\,(1-p)\,\frac{1-\left(\frac{1-p}{p}\right)^{k-1}}{1-\frac{1-p}{p}} \\
&= \frac{p^k(1-p) - (1-p)^k p}{2p-1}
\end{aligned}
$$

To calculate EY, we first note

$$
\sum_{k=2}^{\infty} k\,p^k(1-p) = p^2\sum_{k=2}^{\infty} k\,p^{k-2}(1-p) = p^2\sum_{k'=1}^{\infty}(k'+1)\,p^{k'-1}(1-p) = p^2\left(\frac{1}{1-p}+1\right)
$$

Similarly

$$
\sum_{k=2}^{\infty} k\,(1-p)^k\,p = (1-p)^2\,\left(\frac{1}{p}+1\right)
$$

Thus we get

$$
EY = \frac{1}{2p-1}\,\left(p^2\left(\frac{1}{1-p}+1\right) - (1-p)^2\,\left(\frac{1}{p}+1\right)\right)
$$

$$
\begin{aligned}
&= \frac{1}{2p-1}\left(\frac{p^3-(1-p)^3}{p(1-p)} + \left(p^2-(1-p)^2\right)\right) \\
&= \frac{1}{2p-1}\left(\frac{(p-(1-p))(p^2+p(1-p)+(1-p)^2)}{p(1-p)} + (2p-1)\right) \\
&= \frac{(p+(1-p))^2-p(1-p)}{p(1-p)} + 1 \\
&= \frac{1}{p(1-p)}
\end{aligned}
$$

b. For any two random variables, X, Y, show that $\mathrm{Cov}(X,Y) = \mathrm{Cov}(X, E[Y|X])$

Answer:

$$
\begin{aligned}
\mathrm{Cov}(X, E[Y|X]) &= E\left[X E[Y|X]\right] - EX\ E[E[Y|X]] \\
&= E\left[\ E[XY|X]\ \right] - EX\ EY = E[XY] - EX\ EY = \mathrm{Cov}(X,Y)
\end{aligned}
$$

Marks for part (a) – 7
Marks for part (b) – 3

4. a. Let X be a discrete random variable taking non-negative integer values with mass function, $p(i),\ i = 0, 1, \cdots$. Let $Y_1, Y_2, \cdots, Y_n$ be *iid* discrete random variables taking non-negative integer values and with mass function $q(i), i = 0, 1, \cdots$. Assume $p(i), q(i) > 0, \forall i$. Let $h : \Re \to \Re$ be some function. Define

$$
S = \frac{1}{n}\sum_{k=1}^{n} \frac{p(Y_k)h(Y_k)}{q(Y_k)}.
$$

Find ES.

Answer: For any function g, $E[g(Y_k)] = \sum_m g(m)\ q(m)$. Hence

$$
E\left[\frac{p(Y_k)h(Y_k)}{q(Y_k)}\right] = \sum_{m=0}^{\infty} \frac{p(m)h(m)}{q(m)}\ q(m) = \sum_{m=0}^{\infty} p(m)h(m) = E[h(X)]
$$

Hence we get

$$
ES = \frac{1}{n}\sum_{k=1}^{n}\ E\left[\frac{p(Y_k)h(Y_k)}{q(Y_k)}\right] = \frac{1}{n}\sum_{k=1}^{n} E[h(X)] = E[h(X)]
$$

Comment: Suppose $X_1, X_2, \cdots, X_n$ are iid with mass function $p(i)$. That is, they are iid realizations or iid samples of X. Then we know that $S = \frac{1}{n}\sum_{i=1}^{n} h(X_i)$ is an unbiased estimator of $E[h(X)]$. (That is, $ES = E[h(X)]$). But suppose generating sample of X is difficult. Though we know the expression for $p(i)$, it may be computationally costly to generate samples from there. Suppose $q(i)$ is another distribution from which we can easily generate samples and suppose $Y_1, \cdots, Y_n$ are the samples. We want to use the 'average' of $h(Y_i)$ to estimate $E[h(X)]$. What the above says is that we can do so but then the average is not taken with same weight to all samples. The weight we give to $h(Y_k)$ in this average is equal to $\frac{p(Y_k)}{q(Y_k)}$. If we weight different samples like this then we get an estimate of $E[h(X)]$. This problems illustrates a special case of a general technique known as importance sampling.

b. Let X, Y be jointly Gaussian with means zero, variances 1 and correlation coefficient ρ. Assume $\rho \neq 0$. Let $Z = aX + bY$ and $W = bX + aY$, where $a, b \in \Re$, $a \neq 0, b \neq 0$. Find a sufficient condition on a, b for Z and W to be independent.

Answer: We are given

$$\begin{bmatrix} Z \\ W \end{bmatrix} = \begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Hence, if $a^2 - b^2 \neq 0$ then Z, W are jointly Gaussian. Then they would be independent if they are uncorrelated.

Since $EX = EY = 0$, we have $EZ = EW = 0$. hence

$$\text{Cov}(Z, W) = E[ZW] = E[abX^2 + abY^2 + XY(a^2+b^2)] = 2ab + \rho(a^2+b^2)$$

Hence, Z, W are uncorrelated and hence independent if

$$a^2 - b^2 \neq 0, \quad \text{and} \quad \frac{-2ab}{a^2+b^2} = \rho$$

Comment: Since X, Y are jointly Gaussian, you know there is a linear transform to make them independent. For example, the transformation that would diagonalize the covariance matrix. But that is not what is required here. For the particular form of linear transform given, you are asked to find conditions on a, b to make Z, W independent.

Marks for part (a) – 5
Marks for part (b) – 5

E1 222 Stochastic Models and Applications
Test III

`Time: 75 minutes` `Max. Marks:40`
`Date: 13 Jan 2021`

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

1. a. Let $X_1, X_2, \cdots$ be a sequence of iid continuous random variables. Let their common distribution function be F and suppose it is strictly monotonically increasing. Let $M_n = \max(X_1, \cdots, X_n)$ and $Y_n = n[1 - F(M_n)],\ n = 1, 2, \cdots$. Find the limiting distribution of Y_n.

Answer: Given $M_n = \max(X_1, \cdots, X_n)$. Hence distribution of M_n is

$$F_{M_n}(z) = P[M_n \leq z] = P[X_i \leq z,\ 1 = 1, \cdots, n] = (F(z))^n$$

Since we are given that F is strictly increasing and X_i are continuous rv, we know F is invertible. Now the distribution of Y_n is

$$\begin{aligned} F_{Y_n}(y) &= P[n[1 - F(M_n)] \leq y] \\ &= P\left[F(M_n) \geq 1 - \frac{y}{n}\right] \\ &= P\left[M_n \geq F^{-1}\left(1 - \frac{y}{n}\right)\right] \\ &= 1 - P\left[M_n \leq F^{-1}\left(1 - \frac{y}{n}\right)\right] \\ &= 1 - \left(F\left(F^{-1}\left(1 - \frac{y}{n}\right)\right)\right)^n \\ &= 1 - \left(1 - \frac{y}{n}\right)^n \end{aligned}$$

Hence

$$\lim_{n\to\infty} F_{Y_n}(y) = 1 - \lim_{n\to\infty}\left(1 - \frac{y}{n}\right)^n = 1 - e^{-y}$$

Hence limiting distribution of Y_n is exponential.

b. Let $X_1, X_2, \cdots$ be iid continuous random variables with density $f(x) = 12x^2(1-x),\ 0 \leq x \leq 1$. Let $S_n = \sum_{i=1}^n X_i^2$. Does $\frac{1}{n}S_n$ converge almost surely? Answer Yes/No with a short justification. If your answer is yes, find the limit.

Answer: Yes. Since X_i are iid, so are X_i^2 and, hence, by strong law of large numbers, $\frac{1}{n}S_n$ converges to $E[X_i^2]$.

$$E\left[X_i^2\right] = \int_0^1 x^2\ 12x^2(1-x)\ dx = 12\left[\frac{x^5}{5}\bigg|_0^1 - \frac{x^6}{6}\bigg|_0^1\right] = \frac{12}{30} = \frac{2}{5}$$

So, $\frac{1}{n}S_n$ converges to $\frac{2}{5}$

2. a. Consider a Probability space $(\Omega, \mathcal{F}, P)$ where $\Omega = \{1, 2, \cdots\}$, $\mathcal{F}$ is the power set of Ω and $P(\{i\}) = q_i,\ \forall i$. Note that we would have $q_i \geq 0, \forall i$ and $\sum_i q_i = 1$. Let $X_1, X_2, \cdots$ be a sequence of discrete random variables defined on this space given by

$$\begin{aligned} X_n(\omega) &= 1 \ \text{ if } \ n \leq \omega \\ &= 0 \ \text{ otherwise} \end{aligned}$$

Does the sequence converge in (i) Probability, (ii) almost surely.

Answer: By the definition of X_n, we have

$$P[X_n = 1] = P\left(\{\omega\ :\ \omega \geq n\}\right] = P\left(\{n, n+1, \cdots\}\right) = 1 - \sum_{k=1}^{n-1} q_k$$

This goes to zero as n goes to infinity. We can verify that $X_n \xrightarrow{P} 0$ as:

$$\lim_{n\to\infty} P[|X_n - 0| > \epsilon] = \lim_{n\to\infty} P[X_n = 1] = 1 - \lim_{n\to\infty} \sum_{k=1}^{n-1} q_k = 0$$

We cannot use Borel-Cantelli lemma here for checking for almost sure convergence. Whether or not $\sum_{n=1}^{\infty}\sum_{k=n}^{\infty} q_k < \infty$ depends on q_k.

But since we are given the random variables as functions on Ω, we can ask when does $X_n(\omega) \to 0$. Fix any ω. Then $X_n(\omega)$ is 1 till $n \leq \omega$; after that it stays zero. So, $X_n(\omega) \to 0$ as $n \to \infty$, $\forall \omega$. Hence, X_n converges almost surely to zero

b. A university has 300 vacancies for research students. Since not all students offered admission would accept, the university sends out offers of admission to 400 students. By past experiance the university knows that only 70% of students offered admission would accept the offer. Calculate the approximate probability that more than 300 students would accept the offer of admission.

Answer: Let X_i be indicator random variables representing whether or not the i^{th} student offered admission would accept. We assume that these are independent. Then X_i are iid binary random variables with $[X_i = 1] = 0.7$. Let $S_{400} = \sum_{i=1}^{400} X_i$. S_{400} would be the number of students who have accepted the offer. Hence,

$$
\begin{aligned}
P[S_{400} > 300] &= P\left[\frac{S_{400} - ES_{400}}{\sqrt{\text{Var}(S_{400})}} > \frac{300 - ES_{400}}{\sqrt{\text{Var}(S_{400})}}\right] \\
&= P\left[\frac{S_{400} - ES_{400}}{\sqrt{\text{Var}(S_{400})}} > \frac{300 - 400 * 0.7}{\sqrt{400 * 0.7 * 0.3}}\right] \\
&= P\left[\frac{S_{400} - ES_{400}}{\sqrt{\text{Var}(S_{400})}} > \frac{1}{\sqrt{0.7 * 0.3}}\right] \\
&= P\left[\frac{S_{400} - ES_{400}}{\sqrt{\text{Var}(S_{400})}} > 2.18\right] \\
&\approx 1 - \Phi(2.18) = 1 - 0.985 = 0.015
\end{aligned}
$$

Comment: I did not use the so called continuity correction in the above. We could have, for example, calculated $P[S_n > 300.5] \approx 0.013$ as the required probability.

3. a. Consider an irreducible birth-death Markov chain on the state space $\{0, 1, \cdots, N\}$. If the chain is started in state 1 what is the probability that the chain will visit state $N - 1$ at sometime or the other. Can this chain have a null recurrent state? Explain your answer.

Answer: Since this is a finite irreducible chain, all states are recurrent. Hence, starting from any state the probability of visiting any other state at some finite time is 1.

Since this is a finite chain it cannot have any null recurrent states. This can be established as follows. We can prove that (i). any finite closed set has to have at least one positive recurrent state and (ii). if x is positive recurrent and x leads to y then y is positive recurrent.

In a finite chain all sets of closed irreducible sets of recurrent states would be finite. Since it is a finite closed set, it should have at least one positive recurrent state and since it is irreducible, now all states in the set have to be positive recurrent. Thus each of the finite sets closed irreducibles sets of recurrent states have to be wholly positive recurrent and hence we cannot have a null recurrent state.

Comment: I am only looking for some logical explanation such as the one above. I am not looking for a formal prrof of the two results listed above. Since this is an open-notes exam, I assumed it should be clear to you that I would not be asking you to copy proofs from your notes.

b. Consider a Markov chain with the following transition probability matrix:

$$P = \begin{bmatrix} 0.15 & 0.22 & 0.1 & 0.28 & 0.25 \\ 0 & 0.25 & 0.75 & 0 & 0 \\ 0 & 0.75 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0.65 & 0.35 \\ 0 & 0 & 0 & 0.55 & 0.45 \end{bmatrix}$$

Specify which are the transient and recurrent states and find all the closed irreducible subsets of recurrent states. Find a stationary distribution of the chain

Answer: Let us label the state space as $\{0, 1, \cdots, 4\}$. From the transition probability matrix, it is easy to see that the set $\{1, 2\}$ is closed and irreducible. Since this is a finite closed set, there has to be at least one recurrent state and since it is irreducible, both states in the set are recurrent. Since state 0 leads to this closed set, we can conclude 0 is transient. From the matrix we can also similarly see that the set $\{3, 4\}$ is the other closed and irreducible set of recurrent states. Thus we get

$$S_T = \{0\}, \quad S_R = \{1, 2\} + \{3, 4\}$$

For a chain like this there is always a stationary distribution concentrated on any one set of closed and irreducible set. If we pick the set $\{1, 2\}$, the transition probability matrix for this subset is doubly stochastic and hence we know what is the unique stationray distribution if we consider only this subset. Thus a stationary distribution for the chain is $\pi = [0 \;\; 0.5 \;\; 0.5 \;\; 0 \;\; 0]$

4. a. A man has n umbrellas. Everyday in the morning he goes from his house to office and takes an umbrella with him if it is raining and if he has an umbrella with him; he goes without an umbrella if it is not raining or if he has no umbrellas with him. Similarly in the evening when he goes from office to home he takes an umbrella if it is raining and he has one. The probability of rain is same in the morning and evening and it is equal to p. Construct an $n+1$ state Markov chain and using that calculate the probability that the man would be without an umbrella when it is raining. (Note that this is the generalization of the problem solved in class)

Answer: As discussed in the class a useful way to formulate this is to take the number of umbrellas with the man as the state. Thus we have the state space $\{0, 1, \cdots, n\}$.

From state 0, you can only go to n. Consider a state i, $i \geq 1$. That means the man has i umbrellas with him. So, if it is not raining the state would change to $n - i$ because he would not carry an umbrell; if it is raining the state would change to $n-i+1$ because he would carry an umbrella. Thus the transition probabilities are given by

$$P(0, n) = 1; \quad P(i, n-i) = 1-p, \;\; P(i, n-i+1) = p, \; i = 1, \cdots, n$$

It is easily seen that the chain is irreducible. This can be seen as follows. Suppose you want to go from state i to j and assume $i < j$. You make a transition out of i with probability $1-p$ and from the next state make a transition with probability p. Then you would be in state $i+1$. That is, currently the man has i umbrellas, if he goes to the other place without carrying an umbrella and comes back carrying an umbrella, then he would have $i+1$ umbrellas with him. Like this you can go from i to j. Similar arguments apply when $j < i$

Since this is a finite irreducible chain it has a unique stationary distribution. The stationary distribution has to satisfy

$$\pi(j) = \sum_i P(i, j)\pi(i)$$

For a given $j \neq 0$, $P(i, j) \neq 0$ only if $i = n-j$ or $i = n-j+1$. You can come to state 0 only from n Thus the stationary distribution has to satisfy

$$\begin{aligned}
\pi(0) &= (1-p)\pi(n), \quad \text{and} \\
\pi(j) &= P(n-j,j)\pi(n-j) + P(n-j+1,j)\pi(n-j+1) \\
&= (1-p)\pi(n-j) + p\pi(n-j+1), \; j = 1, \cdots, n-1 \\
\pi(n) &= \pi(0) + p\pi(1)
\end{aligned}$$

It is easy to see that all these equations would be satisfied if choose $\pi(i) = \alpha, \; i = 1, \cdots, n$ and $\pi(0) = (1-p)\alpha$. Then we need $n\alpha + (1-p)\alpha = 1$ which gives $\alpha = \frac{1}{n+1-p}$. Thus the stationary distribution is

$$\pi(j) = \frac{1}{n+1-p}, \; j = 1, \cdots, n, \quad \pi(0) = \frac{1-p}{n-1-p}$$

The probability that the man would be without an umbrella when it is raining is $p\pi(0)$.

b. Let $X_n, \; n = 1, 2, \cdots$ be discrete random variables taking values in $\{0, 1, 2, \cdots, K\}, \; K < \infty$. Suppose $X_n \xrightarrow{P} 0$. Then show that the sequence converges in r^{th} mean to zero.

Answer: Since we are given $X_n \xrightarrow{P} 0$, we have

$$\lim_{n \to \infty} P[X_n \neq 0] = 0$$

Now we have

$$0 \leq E\left[|X_n - 0|^r\right] = \sum_{j=1}^{K} j^r P[X_n = j] \leq K^r \sum_{j=1}^{K} P[X_n = j] = K^r P[X_n \neq 0]$$

Since $\lim_{n\to\infty} P[X_n \neq 0] = 0$, $\lim_{n\to\infty} E\left[|X_n - 0|^r\right] = 0$ which proves that X_n converge in r^{th} mean.

Comment: I hope it is easy to see that the restriction that all X_i take values in the same set is not needed. All we need is that there is a $K < \infty$ such that $|X_n| < K$ for all n.

In each question, each part would be graded 5 marks

E1 222 Stochastic Models and Applications
Final Examination

Time: 180 minutes Max. Marks:50
Date: 28 Nov 2019

Answer **any FIVE** questions. All questions carry equal marks
Answers should be written only in the space provided.

1. a. Let A, B be events with $P(A) = 0.3$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$. Let I_A and I_B be the indicator random variables of events A and B respectively. Find the correlation coefficient of I_A and I_B.

   b. Let X be an integer valued random variable with distribution function F. Let Y be uniformly distributed over $(0, 1)$. Define the integer valued random variable Z by

$$Z = m \text{ if } F(m-1) < Y \leq F(m)$$

   Find the distribution of Z.

2. a. Let X be uniform over $[-1, 1]$. Let $Y = \sqrt{|X|}$. Find density of Y and EY.

   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 EY. (The mass function of geometric random variable is: $f(x) = p(1-p)^{x-1},\ x = 1, 2, \cdots$).

3. a. Let X, Y have joint density given by

$$f_{XY}(x, y) = \sqrt{\frac{2}{\pi}}\, e^{-0.5\left((x-y)^2+4y\right)},\ -\infty < x < \infty, 0 < y < \infty$$

   Find $f_Y(y)$, $f_{X|Y}(x|y)$, and $E[X|Y]$.

   b. Let X be a continuous random variable which is uniform over $(0, 1)$. Let Y be a discrete random variable taking non-negative integer values. Assume X and Y are independent. Let $Z = X+Y$. Show that Z is a continuous random variable.

4. a. Let X, Y be two random variables. The conditional variance of X given Y is defined by

$$\mathrm{Var}(X|Y) = E\left[(X - E[X|Y])^2 \,|Y\right]$$

Show that

$$\mathrm{Var}(X) = E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y])$$

b. Let X be a geometric random variable. Calculate EX^3.

5. a. Let $X_1, X_2, \cdots$ be iid continuous random variables with common distribution function F which is strictly monotonically increasing. Define $Y_n = \max(X_1, \cdots, X_n)$. Define $Z_n = n(1 - F(Y_n))$. Does the sequence Z_n converge in distribution?

b. Consider a probability space with $\Omega = \{1, 2, \cdots\}$ and with the set of events being the power set of Ω. Consider the probability assignment $P[\{i\}] = q_i,\ i = 1, 2, \cdots$. (Note that the q_i satisfy: $q_i \geq 0,\ \sum_i q_i = 1$). Consider a sequence of random variables defined on this space given by

$$\begin{aligned} X_n(\omega) &= 1 \text{ if } n > \omega \\ &= 0 \text{ otherwise} \end{aligned}$$

Does this sequence converge in (i) probability, (ii). almost surely

6. a. We have two boxes labelled 1 and 2. We also have d balls numbered $1, 2, \cdots, d$. Initially, some of the balls are kept in box-1 and the rest in box-2. At each instant, an integer from the set $\{1, 2, \cdots, d\}$ is selected at random and the ball labelled by that integer is removed from its current box. Then, one of the boxes is selected at random and this ball is put in that box. Let X_n represent the number of balls in box-1 after the operation at time instant n. (X_0 is the initial number of balls in box-1). Provide a brief argument to say that $\{X_n\}$ is a Markov chain. Find its transition probability matrix. For the case, $d = 3$ find the stationary probabilities of the chain.

b. Let $\{X_n,\ n \geq 0\}$ be a Markov Chain. Let s_0, s_1, s_2 be some specific three states. Suppose the probabilities of transition out of

s_0 are given by: $P(s_0, s_0) = 0.25; P(s_0, s_1) = 0.4; P(s_0, s_2) = 0.35$. Suppose the chain is started in s_0. Let T denote the first time instant when the chain left state s_0. (That is, $T = \min\{n : n \geq 1,\ X_n \neq s_0\}$). Find the distribution of T and X_T.

7. a. Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ. Let X_0 be a discrete random variable that is independent of $N(t)$ and with mass function $P[X_0 = +1] = P[X_0 = -1] = 0.5$. Define a stochastic process: $X(t) = X_0\ (-1)^{N(t)}$. Find the mean and autocorrelation function of $X(t)$. Is this process wide-sense stationary?

   b. Vehicles passing a certain point on a highway is a poisson process with rate one per two minutes. Assume that 5% of the vehicles on the highway are trucks and the rest are cars. (i). What is the expected number of trucks that pass the point in one hour? (ii). Given that 8 trucks passed the point in one hour, what is the expected number of cars in that one hour?

E1 222 Stochastic Models and Applications
Final Examination

`Time: 3 hours` `Max. Marks:50`
`Date: 21 Jan 2021`

Answer **Any FIVE** questions. All questions carry equal marks

1. a. Let A, B be events with $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.3$. Let I_A and I_B be the indicator random variables of events A and B respectively. Find the correlation coefficient of I_A and I_B.

Answer: We have

$$E[I_A] = P[I_A = 1] = P(A) = 0.4, \quad E[I_B] = P(B) = 0.5$$

and

$$E[I_A \, I_B] = P[I_A = 1, I_B = 1] = P(A \cap B) = 0.3$$

Hence,

$$\mathrm{Cov}(I_A, I_B) = E[I_A I_B] - E[I_A] \, E[I_B] = 0.3 - 0.4 * 0.5 = 0.1$$

Also, $\mathrm{Var}(I_A) = 0.4 * 0.6 = 0.24$ and $\mathrm{Var}(I_B) = 0.5 * 0.5 = 0.25$. Hence

$$\rho_{I_A I_B} = \frac{0.1}{\sqrt{0.24 * 0.25}} \approx 0.41$$

b. A rod of length 1 is broken at a random point. The piece containing the left end is once again broken at a random point. Let L be the length of the final piece containing the left end. Find the probability that L is greater than 0.25.

Answer: Let us take the rod as the interval [0, 1]. Let X_1 denote the first point where the rod is broken and let X_2 denote the second point. Note that $L = X_2$. We are given that X_1 is uniform from 0 to 1 and X_2 is uniform from 0 to X_1. Thus

$$f_{X_1}(x_1) = 1, \ 0 \le x_1 \le 1 \quad \text{and} \quad f_{X_2|X_1}(x_2|x_1) = \frac{1}{x_1}, \ 0 \le x_2 \le x_1 \le 1$$

This gives us

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_2|X_1}(x_2|x_1) \, f_{X_1}(x_1) \, dx_1 = \int_{x_2}^{1} \frac{1}{x_1} \, dx_1 = -\ln(x_2), \ 0 \le x_2 \le 1$$

Hence

$$P[L > 0.25] = \int_{0.25}^{1} (-\ln(x_2))\ dx_2 = -\ x_2\ \ln(x_2)|_{0.25}^{1} + \int_{0.25}^{1} dx_2 = 0.25*\ln(0.25) + 0.75 \approx 0.4$$

2. a. Consider a game where N men put all their hats in a heap and then everyone randomly chooses a hat. Let$X$ denote the number of men who get their own hat. Show that $\text{Var}(X) = 1$.

Answer: Let $Y_i,\ i = 1, 2, \cdots, N$ be defined by

$$Y_i = \begin{cases} 1 & \text{if } i^{th} \text{ man gets his own hat} \\ 0 & \text{Otherwise} \end{cases}$$

Now we have $X = \sum_{i=1}^{N} Y_i$. Hence

$$\text{Var}(X) = \sum_{i=1}^{N} \text{Var}(Y_i) + \sum_{i=1}^{N} \sum_{j \neq i} \text{Cov}(Y_i, Y_j)$$

We have

$$P[Y_i = 1] = \frac{(N-1)!}{N!} = \frac{1}{N} \quad \Rightarrow \quad E[Y_i] = \frac{1}{N},\ \ \text{Var}(Y_i) = \frac{1}{N}\left(1 - \frac{1}{N}\right)$$

For $i \neq j$,

$$P[Y_i = 1, Y_j = 1] = \frac{(N-2)!}{N!} = \frac{1}{N(N-1)} \quad \Rightarrow \quad E\left[Y_i\ Y_j\right] = \frac{1}{N(N-1)} \quad \Rightarrow \quad \text{Cov}(Y_i, Y_j) = \frac{1}{N(N-1)} - \frac{1}{N^2}$$

Hence, we get variance of X as

$$\begin{aligned} \text{Var}(X) &= \sum_{i=1}^{N} \frac{1}{N}\left(1 - \frac{1}{N}\right) + \sum_{i=1}^{N} \sum_{j \neq i} \left(\frac{1}{N(N-1)} - \frac{1}{N^2}\right) \\ &= \left(1 - \frac{1}{N}\right) + 1 - \frac{N-1}{N} \\ &= 1 \end{aligned}$$

b. Let X, Y be iid exponential random variables with parameter λ. Find the density of $Z = Y/X$

Answer: Let $Z = Y/X$ and $W = X$. This is an invertible transformation: $X = W$ and $Y = ZW$. The Jacobian of the inverse transformation is $\begin{vmatrix} 0 & 1 \\ w & z \end{vmatrix} = -w$. Hence we have

$$f_{ZW}(z, w) = |w| f_{XY}(w, zw)$$

using this we get

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(w, zw) \; dw$$

We could have also used this formula without explicitly deriving it here.

Since X, Y are exponential, the integrand in the above integral is zero unless we have $w \geq 0$ and $zw \geq 0$. Hence we must have $z \geq 0$ and w ranges from 0 to ∞.

$$\begin{aligned}
f_Z(z) &= \int_0^{\infty} w \; \lambda e^{-\lambda w} \; \lambda e^{-\lambda z w} \; dw \\
&= \int_0^{\infty} w \; \lambda^2 \; e^{-\lambda(1+z)w} \; dw \\
&= \frac{\lambda}{1+z} \int_0^{\infty} w \; \lambda(1+z) \; e^{-\lambda(1+z)w} \; dw \\
&= \frac{\lambda}{1+z} \; \frac{1}{\lambda(1+z)} \\
& \quad \text{using formula for expectation of exponential rv} \\
&= \frac{1}{(1+z)^2}
\end{aligned}$$

Hence the density of Z is given by

$$f_Z(z) = \frac{1}{(1+z)^2}, \; 0 \leq z < \infty$$

3. a. Let $X_1, \cdots, X_n$ be *iid* Poisson random variables with mean 1. Let $S_n = \sum_{k=1}^{n} X_k$. Find Prob$[S_n \leq n]$. Show that

$$\lim_{n \to \infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \; = \; 0.5.$$

(You can use the fact: if X, Y are independent Poisson random variables then $X + Y$ is also Poisson).

Answer: If X and Y are independent Poisson with means λ_1 and λ_2 then $X+Y$ is Poisson with mean $\lambda_1+\lambda_2$.

Since each of the X_i are Poisson with mean 1, S_n is Poisson with mean n. So, we get

$$P[S_n \le n] = \sum_{k=0}^{n} \frac{n^k}{k!}\, e^{-n}$$

Now consider a sequence of iid random variable, $X_1, \cdots$, which are all poisson with mean 1. Let $S_n = \sum_{k=1}^{n} X_k,\ n = 1, 2, \cdots$ Since S_n is Poisson with parameter n, we know $ES_n = n$ and $\mathrm{Var}(S_n) = n$. Now, by central limit theorem we get

$$\lim_{n\to\infty} P[S_n \le n] = \lim_{n\to\infty} P\left[\frac{S_n - ES_n}{\sqrt{\mathrm{Var}(S_n)}} \le \frac{n-n}{\sqrt{n}}\right] = \lim_{n\to\infty} P\left[\frac{S_n - ES_n}{\sqrt{\mathrm{Var}(S_n)}} \le 0\right] = \Phi(0) = 0.5$$

This shows that

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \ = \ 0.5.$$

Comment: Please note that central limit theorem does not allow you to conclude that $P[S_n \le n] = 0.5$. It only allows you to conclude $\lim_{n\to\infty} P[S_n \le n] = 0.5$

b. Let $X_1, X_2, \cdots, X_n$ be iid continuous random variables each having uniform distribution over $(0, 1)$. Let $Y_1 = X_1$, $Y_2 = X_1X_2$, $Y_3 = X_1X_2X_3$, $\cdots$, $Y_n = X_1X_2\cdots X_n$. Find joint density of $Y_1, Y_2, \cdots, Y_n$, and conditional density of Y_k conditioned on $Y_1, Y_2, \cdots, Y_{k-1}$. Let t be a fixed number in the interval $[0,\ 1]$. Let Z denote the number of Y_i that are in the interval $[t, 1]$. Find $P[Z = 1]$.

Answer: The given transformation is

$$\begin{aligned}
Y_1 &= X_1 \\
Y_2 &= X_1\, X_2 \\
Y_3 &= X_1X_2X_3 \\
&\vdots \\
Y_n &= X_1\, X_2\, \cdots\, X_n
\end{aligned}$$

The inverse transformation is

$$\begin{aligned} X_1 &= Y_1 \\ X_2 &= Y_2/Y_1 \\ X_3 &= Y_3/Y_2 \\ &\vdots \\ X_n &= Y_n/Y_{n-1} \end{aligned}$$

The Jacobian of the inverse transformation is

$$\left| \begin{array}{cccccc} 1 & 0 & 0 & 0 & \cdots & 0 \\ -y_2/y_1^2 & 1/y_1 & 0 & 0 & \cdots & 0 \\ 0 & -y_3/y_2^2 & 1/y_2 & 0 & \cdots & 0 \\ & & & \vdots & & \\ 0 & 0 & \cdots & 0 & -y_n/y_{n-1}^2 & 1/y_{n-1} \end{array} \right| = \frac{1}{y_1 y_2 \cdots y_{n-1}}$$

Thus, we get

$$f_{Y_1 Y_2 \cdots Y_n}(y_1, y_2, \cdots, y_n) = \left| \frac{1}{y_1 y_2 \cdots y_{n-1}} \right| f_{X_1 X_2 \cdots X_n}(y_1,\ y_2/y_1,\ y_3/y_2,\ \cdots,\ y_n/y_{n-1})$$

Since $X_1, \cdots, X_n$ are iid uniform over $(0,1)$, their joint density function is 1 if all arguments are between 0 and 1 and is zero otherwise. So, in the above expression, for the joint density of $Y_1, \cdots, Y_n$ to be non-zero, we need

$$0 < y_1 < 1, \quad 0 < y_2/y_1 < 1 \ \Rightarrow\ 0 < y_2 < y_1, \quad 0 < y_3/y_2 < 1 \ \Rightarrow\ 0 < y_3 < y_2, \cdots$$

Thus we get the joint density as

$$f_{Y_1 Y_2 \cdots Y_n}(y_1, y_2, \cdots, y_n) = \frac{1}{y_1 y_2 \cdots y_{n-1}}, \ \ 0 < y_n < y_{n-1} < \cdots < y_2 < y_1 < 1$$

From the above derivation, it is easy to see that for any $k > 1$,

$$f_{Y_1 \cdots Y_k}(y_1, \cdots, y_k) = \frac{1}{y_1 \cdots y_{k-1}}, \ \ 0 < y_k < \cdots < y_1 < 1$$

So, we can calculate the required conditional density as

$$\begin{aligned} f_{Y_k | Y_1 \cdots Y_{k-1}}(y_k \mid y_1, \cdots y_{k-1}) &= \frac{f_{Y_1 \cdots Y_k}(y_1, \cdots, y_k)}{f_{Y_1 \cdots Y_{k-1}}(y_1, \cdots, y_{k-1})} \\ &= \frac{1}{y_{k-1}}, \ \ 0 < y_k < y_{k-1} < \cdots < y_1 < 1 \end{aligned}$$

For the last part of the question, it is given that Z is the number of Y_i that are above t. We know that $Y_1 > Y_2 > \cdots > Y_n$. Hence the event $[Z = 1]$ is same as the event $[Y_1 \geq t \text{ and } Y_2 < t]$.

$$P[Z = 1] = \int_t^1 \int_0^t f_{Y_1 Y_2}(y_1, y_2)\, dy_2\, dy_1 = \int_t^1 \int_0^t \frac{1}{y_1}\, dy_2\, dy_1 = -t\ln(t)$$

4. a. Let X be Gaussian with mean zero and variance 1. Let Z be a discrete random variable that is independent of X and suppose $\text{Prob}[Z = 1] = \text{Prob}[Z = -1] = 0.5$. Let $Y = ZX$. Find density of Y. Are X, Y uncorrelated? Are X, Y jointly Gaussian?

Answer: The distribution function of Y is given by

$$\begin{aligned}
P[Y \leq y] &= P[ZX \leq y] \\
&= P[ZX \leq y \mid Z = 1]P[Z = 1] + P[ZX \leq y \mid Z = -1]P[Z = -1] \\
&= P[X \leq y \mid Z = 1]P[Z = 1] + P[-X \leq y \mid Z = -1]P[Z = -1] \\
&= P[X \leq y]P[Z = 1] + P[X \geq -y]P[Z = -1] \\
&\qquad \text{since } X, Z \text{ are independent} \\
&= 0.5F_X(y) + 0.5(1 - F_X(-y)) \\
&= 0.5F_X(y) + 0.5F_X(y) \\
&\qquad \text{since } X \sim \mathcal{N}(0, 1), \text{ we have } 1 - F_X(-y) = F_X(y) \\
&= F_X(y)
\end{aligned}$$

Thus Y is also Gaussian with mean zero and variance 1.

We know $EX = EY = 0$. We can calculate $E[XY] = E[ZX^2] = EZ\ E[X^2] = 0$, because Z, X are independent and $EZ = 0$. Hence, $\text{Cov}(X, Y) = EXY - EX\ EY = 0$ showing that X, Y are uncorrelated.

If X, Y are jointly Gaussian, since they are uncorrelated they must be independent. But since Y can be either X or $-X$ only, they are not independent. For example $P[Y > 2 \mid 0 \leq X \leq 1] = 0$ but $P[Y > 2] \neq 0$. Hence, X, Y are not jointly Gaussian.

Just to reiterate, the argument is as follows. We have established that X, Y **are uncorrelated and are not independent**. Hence, they are not jointly Gaussian.

b. Let X, Y have joint density given by

$$f_{XY}(x,y) = \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} x^{a-1}(y-x)^{b-1}e^{-\lambda y},\ 0 < x < y < \infty$$

where $a, b, \lambda > 0$ are parameters. Find $E[X|Y]$.

Answer: We need the conditional density $f_{X|Y}$ for which we need the marginal density f_Y.

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\, dx \\
&= \int_0^y \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} x^{a-1}(y-x)^{b-1}e^{-\lambda y}\, dx,\ y > 0 \\
& \quad \text{change the variable: } \ z = x/y \quad \text{we get } dx = y dz, \quad \text{limits become 0 to 1} \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\ e^{-\lambda y} \int_0^1 (yz)^{a-1}\ (y - yz)^{b-1}\ y\ dz \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\ e^{-\lambda y}\ y^{a+b-1} \int_0^1 z^{a-1}\ (1-z)^{b-1}\ dz \\
&= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\ e^{-\lambda y}\ y^{a+b-1}\ \frac{\Gamma(a)\ \Gamma(b)}{\Gamma(a+b)} \\
& \quad \text{by using the beta density: } \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)} \int_0^1 z^{a-1}\ (1-z)^{b-1}\ dz = 1 \\
&= \lambda^{a+b}\ y^{a+b-1}\ e^{-\lambda y}\ \frac{1}{\Gamma(a+b)},\ y > 0
\end{aligned}$$

Thus Y has gamma density with parameters $a + b$ and λ

Now, the conditional density is given by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)}\ \frac{x^{a-1}(y-x)^{b-1}}{y^{a+b-1}},\ 0 < x < y < \infty$$

The conditional expectation is given by

$$\begin{aligned}
E[X \mid Y = y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\, dx \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)} \int_0^y x\ \frac{x^{a-1}(y-x)^{b-1}}{y^{a+b-1}}\ dx \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\ \Gamma(b)} \int_0^y \frac{x}{y} \left(\frac{x}{y}\right)^{a-1} \left(1 - \frac{x}{y}\right)^{b-1}\ dx
\end{aligned}$$

$$
\begin{aligned}
& \text{change variable: } z = \frac{x}{y} \\
= \;& y \, \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \int_0^1 z \, z^{a-1} \, (1-z)^{b-1} \, dz \\
= \;& \frac{a}{a+b} \, y
\end{aligned}
$$

where we have used the formula for mean of beta density.
Thus, $E[X \mid Y] = \frac{a}{a+b} Y$

5. a. Let X_n, $n = 1, 2, \cdots$, be *iid* random variables uniform over $[0,\ 1]$. Let $W_n = \min(X_1, \cdots, X_n)$ and $Y_n = \max(X_1, \cdots, X_n)$. Let $Z_n = 0.5(W_n + Y_n)$. Does the sequence Z_n converge in probability?

Answer: X_n are iid uniform over $[0, 1]$. Given $Y_n = \max(X_1, \cdots, X_n)$. Then $Y_n \xrightarrow{P} 1$. This can be seen as

$$P[|1-Y_n| > \epsilon] = P[Y_n < 1-\epsilon] = P[X_i < 1-\epsilon,\ i = 1, \cdots, n] = (1-\epsilon)^n$$

which goes to zero as $n \to \infty$ for all $\epsilon > 0$.

Similarly, $W_n = \min(X_1, \cdots, X_n)$ converges to zero in probability:

$$P[|W_n - 0| > \epsilon] = P[W_n > \epsilon] = P[X_i > \epsilon,\ i = 1, \cdots, n] = (1-\epsilon)^n$$

which goes to zero as $n \to \infty$ for all $\epsilon > 0$.

Since $W_n \xrightarrow{P} 0$ and $Y_n \xrightarrow{P} 1$, $(W_n + Y_n) \xrightarrow{P} 0 + 1 = 1$. Hence $Z_n = 0.5(W_n + Y_n)$ converges in probability to 0.5.

Comment: While correcting the paper I will accept the answer if you simply state that $(W_n + Y_n) \xrightarrow{P} 0 + 1 = 1$.

As we mentioned in class, $X_n \xrightarrow{P} c_1$ and $Y_n \xrightarrow{P} c_2$ implies $(X_n + Y_n) \xrightarrow{P} c_1 + c_2$. This result can be easily proved as follows.

By definition, $X_n \xrightarrow{P} c$ is same as $(X_n - c) \xrightarrow{P} 0$. Hence, without loss of generality, we can show the result assuming $c_1 = c_2 = 0$.

The result follows because we have

$$P[\,|X_n + Y_n| > \epsilon\,] \le P[\,|X_n| > \epsilon/2\,] + P[\,|Y_n| > \epsilon/2\,]$$

This follows because for $X_n + Y_n$ to be greater than ϵ at least one of X_n or Y_n have to be greater than $\epsilon/2$; if both are less that $\epsilon/2$ their sum cannot exceed ϵ. Since some of you may be new to such arguments (which are common in Analysis), here is a more detailed derivation of this.

Since $|x+y| \leq |x|+|y|$, we have the following, for any two random variables X, Y:

$$([\,|X| \leq \epsilon/2\,] \;\cap\; [\,|Y| \leq \epsilon/2\,]) \;\subset [\,|X+Y| \leq \epsilon\,]$$

Please note that the above is a relation among sets. The above follows because if $|X(\omega)| \leq \epsilon/2$ and $|Y(\omega)| \leq \epsilon/2$, then we have $|X(\omega) + Y(\omega)| \leq |X(\omega)| + |Y(\omega)| \leq \epsilon/2 + \epsilon/2 = \epsilon$.

If $A \subset B$ then, $B^c \subset A^c$. Hence, we get

$$[\,|X+Y| > \epsilon\,] \subset ([\,|X| > \epsilon/2\,] \;\cup\; [\,|Y| > \epsilon/2\,])$$

We know that if $A \subset B$ then $P(A) \leq P(B)$ and also that $P(A \cup B) \leq P(A) + P(B)$.

Hence, now considering the sequneces X_n, Y_n,

$$P[\,|X_n + Y_n| > \epsilon\,] \leq P[\,|X_n| > \epsilon/2\,] + P[\,|Y_n| > \epsilon/2\,]$$

If $X_n \xrightarrow{P} 0$ and $Y_n \xrightarrow{P} 0$, then the RHS above goes to zero as $n \to \infty$ which implies $X_n + Y_n \xrightarrow{P} 0$,

b. Let $X_1, X_2, \cdots$ be iid Gaussian random variables with mean zero and variance 1. Define

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \qquad Z_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

What are the distributions of Y_n, Z_n? Do the sequences Y_n, Z_n converge to some constant in probability? If yes, state the limit. Do these sequences converge in distribution? If yes, state the limit.

(You need not ‘derive’ or ‘prove’ any thing here. Simply state the answer along with a short justification/explanation for the answer).

Answer: We have $EX_i = 0$ and hence $EY_n = EZ_n = 0$. Also

$$\text{Var}(Y_n) = \left(\frac{1}{\sqrt{n}}\right)^2 \sum_{i=1}^{n} \text{Var}(X_i) = 1, \ \forall n, \quad \text{and} \ \ \text{Var}(Z_n) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \text{Var}(X_i) = \frac{1}{n}, \ \forall n$$

Since X_i are iid Gaussian, any linear combination of them would be Gaussian. Hence Y_n would be Gaussian with zero mean and variance 1 for all n. Similarly, Z_n would be Gaussian with mean zero and variance $1/n$.

Please note that the fact that Z_n and Y_n are Gaussian for every n has nothing to do with CLT. This comes about because linear combinations of independent Gaussians are Gaussian. The CLT tells you only about the limit distribution. It cannot tell you what the actual distribution is of a sum of finite number iid random variables. But, CLT can be used to **approximate** this distribution if the number in the sum is large.

Since all Y_n have the same density, the sequence cannot converge to any constant in probability. This is because for a given ϵ and any fixed finite c, $P[|Y_n - c| > \epsilon]$ is a fixed non-zero number independent of n and hence does not go to zero as n tends to infinity. Another way of looking at this is that the variance of Y_n is the same non-zero value for all n.

Since all Y_n have the same distribution, namely, standard Normal, the sequence converges in distribution to that.

Z_n converges to zero in probability by weak law of large numbers.

Since convergence in probability implies convergence in distribution, the sequence Z_n converges in distribution to the degenerate distribution representing the constant 0. (Note that the limit distribution here is: $F(x) = 0$ if $x < 0$ and $F(x) = 1$ if $x \geq 0$. This is the distribution function of a discrete random variable that takes only one value, namely, zero. The variance of this distribution is zero. That is why it is termed the degenerate distribution. See problem 6 in problem sheet 4.1)

6. a. Let $X(t)$ be a stochastic process defined by $X(t) = (-1)^{N(t)}\, X_0$ where $N(t)$ is a Poisson process with rate λ and X_0 is a random variable which is independent of $N(t)$ and which has the distribution $P[X_0 = +1] = P[X_0 = -1] = 0.5$. Find the mean and autocorrelation of $X(t)$.

Answer: The mean of $X(t)$ is, using independence of X_0 and the process $N(t)$ and the fact that $EX_0 = 0$,

$$E[X(t)] = E\left[(-1)^{N(t)}\ X_0\right] = E[X_0]\ E\left[(-1)^{N(t)}\right] = 0$$

The autocorrelation function is given by, for $t, s > 0$,

$$\begin{aligned}
R_X(t, t+s) &= E\left[X_0^2\ (-1)^{N(t)+N(t+s)}\right] \\
&= E\left[(-1)^{N(t)+N(t+s)}\right], \quad \text{since } X_0^2 = 1 \\
&= E\left[(-1)^{N(t)+N(t+s)-N(t)+N(t)}\right] \\
&= E\left[(-1)^{2N(t)}(-1)^{N(t+s)-N(t)}\right] \\
&= E\left[(-1)^{N(t+s)-N(t)}\right] \\
&= E\left[(-1)^{N(s)}\right], \quad \text{since the process has stationary increments} \\
&= \sum_{k=0}^{\infty}(-1)^k\ \frac{(\lambda s)^k}{k!}e^{-\lambda s} \\
&= e^{-\lambda s}\sum_{k=0}^{\infty}\frac{(-\lambda s)^k}{k!} \\
&= e^{-2\lambda s}
\end{aligned}$$

Thus $E[X(t)] = 0$ and $R_X(t, t+s) = e^{-2\lambda s}$.

b. Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ and assume that it is independent of a non-negative random variable, T. Suppose the mean of T is μ and its variance is σ^2. Find (i). $E[N(T)]$, (ii). $\mathrm{Var}(N(T))$

Answer: For the first part, we can use $E[N(T)] = E[\ E[N(T) \mid T]\ ]$

$$E[N(T) \mid T = t] = E[N(t) \mid T = t] = E[N(t)] = \lambda t$$

In the above we have used the independence of T and $N(t)$. Thus, we have $E[N(T) \mid T] = \lambda T$. Hence, $E[N(T)] = E[\lambda T] = \lambda\ \mu$. Similarly,

$$E[N^2(T) \mid T = t] = E[N^2(t) \mid T = t] = E[N^2(t)] = \lambda t + (\lambda t)^2$$

Thus,

$$E[N^2(T) \mid T] = \lambda T + \lambda^2 T^2$$

which gives us

$$E[N^2(T)] = \lambda\, \mu + \lambda^2(\sigma^2 + \mu^2)$$

Hence

$$\text{Var}(N(T)) = E[N^2(T)] - (E[N(T)])^2 = \lambda\mu + \lambda^2(\sigma^2 + \mu^2) - \lambda^2\mu^2 = \lambda\mu + \lambda^2\sigma^2$$

We could have also got the variance of $N(T)$ using the formula

$$\text{Var}(N(T)) = \text{Var}(E[N(T)|T]) + E[\text{Var}(N(T)|T)]$$

7. a. Let $\{X_n,\ n \geq 0\}$ be a Markov Chain. Let s_0, s_1, s_2 be some specific three states. Suppose the probabilities of transition out of s_0 are given by: $P(s_0, s_0) = 0.5; P(s_0, s_1) = 0.2; P(s_0, s_2) = 0.3$. Suppose the chain is started in s_0. Let T denote the first time instant when the chain left state s_0. (That is, $T = \min\{n : n \geq 1,\ X_n \neq s_0\}$). Find the distribution of T and X_T.

Answer: The event $[T = k]$ is same as $[X_1 = s_0, \cdots, X_{k-1} = s_0, X_k \neq s_0]$. Hence

$$P[T = k|X_0 = s_0] = (P(s_0, s_0))^{k-1}(1 - P(s_0, s_0)) = 0.5^{k-1}0.5,\ k = 1, 2, \cdots$$

Thus, T is geometric with parameter 0.5.

X_T is the state of the chain at the instant when the chain has left s_0. Hence $X_T \in \{s_1, s_2\}$.

To get the distribution of X_T we can argue intuitively as follows. Consider the chain started in s_0. We can think of the chain making a transition as a random experiment whose out come is the next state. If we consider the chain only till the first instant when it is not in s_0, for this random experiment $\Omega = \{s_0, s_1, s_2\}$. We can think of the situation as independent repetitions of this experiment till one of the events $\{s_1\}$ or $\{s_2\}$ occurs. The event of $[X_T = s_1]$ is same as $\{s_1\}$ occurring before $\{s_2\}$ in this repetition of our random experiment. Hence we can conclude

$$P[X_T = s_1] = \frac{P(s_0, s_1)}{P(s_0, s_1) + P(s_0, s_2)} = \frac{0.2}{0.2 + 0.3} = 0.4$$

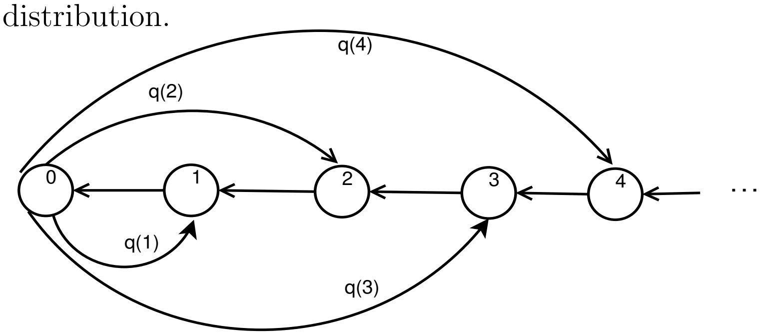
Hence $P[X_T = s_2] = 0.6$.

We can get it more formallylike as follows. Note all probabilities are conditioned on chain starting in s_0 though we do not explicitly show it.

$$\begin{aligned}
P[X_T = s_1] &= \sum_{k=1}^{\infty} P[X_T = s_1, T = k] \\
&= \sum_{k=1}^{\infty} P[X_1 = s_0, \cdots, X_{k-1} = s_0, X_k = s_1] \\
&= \sum_{k=1}^{\infty} (P(s_0, s_0))^{k-1} P(s_0, s_1) \\
&= \frac{P(s_0, s_1)}{1 - P(s_0, s_0)} \\
&= \frac{0.2}{0.2 + 0.3} = 0.4
\end{aligned}$$

And, hence, $P[X_T = s_2] = 0.6$.

b. Consider the following Markov chain on state space $\{0, 1, \cdots\}$. Take $q(k) = (1-p)^{k-1}p,\ k = 1, 2, \cdots$ with $0 < p < 1$. Will this chain have a stationary distribution? If yes, find the stationary distribution.



Answer: This is an irreducible Markov chain. It would have a unique stationary distribution if and only if it is positive recurrent. We can show the chain to be positive recurrent by calculating mean return time to state 0. Alternately, we can directly show that there is a distribution that satisfies the conditions for stationary distribution.

Suppose π is a stationarydistribution. Then it has to satisfy

$$\pi(j) = \sum_i \pi(i)P(i,j)$$

Taking $j = 0, 1, \cdots$, we get

$$\begin{aligned}
\pi(0) &= \pi(1) \\
\pi(1) &= \pi(0)q(1) + \pi(2) \quad \Rightarrow \quad \pi(2) = \pi(1) - q(1)\pi(0) = (1-q(1))\pi(0) \\
& \qquad \text{where we have used } \; \pi(1) = \pi(0) \\
\pi(2) &= \pi(0)q(2) + \pi(3) \quad \Rightarrow \quad \pi(3) = (1 - q(1) - q(2))\,\pi(0) \\
& \qquad \text{where we used } \; \pi(2) = (1-q(1))\pi(0) \\
& \vdots \\
\pi(k-1) &= \pi(0)q(k-1) + \pi(k) \quad \Rightarrow \quad \pi(k) = \left(1 - \sum_{j=1}^{k-1} q(j)\right)\pi(0) \\
& \vdots
\end{aligned}$$

Since $\left(1 - \sum_{j=1}^{k-1} q(j)\right) = \sum_{j=k}^{\infty} q(j)$, we get

$$\pi(k) = \pi(0)\sum_{j=k}^{\infty} q(j) = \pi(0)\sum_{j=k}^{\infty} p(1-p)^{j-1} = \pi(0)(1-p)^{k-1}, \; k = 1, 2, \cdots$$

This would be a stationary distribution if we can find $\pi(0)$ (with $0 < \pi(0) \le 1$) such that $\sum_{j=0}^{\infty} \pi(j) = 1$. Thus we need

$$\pi(0) + \pi(0)\sum_{k=1}^{\infty}(1-p)^{k-1} = \pi(0) + \pi(0)\frac{1}{p} = 1$$

This implies $\pi(0) = \frac{p}{1+p}$. Hence the stationary distribution is

$$\pi(0) = \frac{p}{1+p}, \quad \pi(k) = \frac{p(1-p)^{k-1}}{1+p}, \; k = 1, 2, \cdots$$

Comment: As I said above, we could have actaully calculated mean recurrence time for state 0. Consider the chain started in state 0. Let T be time to return to 0. Then it is easy to see that

$$P[T=1] = 0, \quad \text{and} \quad P[T=k] = q(k-1), \; k = 2, \cdots$$

Hence

$$m_0 = E_0[T] = \sum_{k=1}^{\infty}(k+1)q(k) = \sum_{k=1}^{\infty}(k+1)(1-p)^{k-1}p = \frac{1}{p} + 1$$

Since $m_0 < \infty$, we can conclude that this irreducible chain is positive recurrent. Also, we know $\pi(0) = 1/m_0 = p/(1+p)$ which is same as what we calculated above. However, to get $\pi(1), \pi(2)$, etc. we need to write the equations for a stationary distribution and solve them.

8. a. Let $\{B(t),\ t \geq 0\}$ be a standard Brownian motion process. Consider a process defined by

$$V(t) = e^{-\alpha t/2}\ B\left(e^{\alpha t}\right)$$

where $\alpha > 0$ is a parameter. Find the mean and autocorrelation of $V(t)$.

Answer: The mean of the process is given by

$$E[V(t)] = e^{-\alpha t/2}\ E\left[B\left(e^{\alpha t}\right)\right] = 0$$

because $E\left[B(t)\right] = 0,\ \forall t$.

The autocorrelation function is, for $t, s > 0$

$$\begin{aligned} R(t, t+s) &= e^{-\alpha t/2}\ e^{-\alpha(t+s)/2} E\left[B\left(e^{\alpha t}\right)\ B\left(e^{\alpha(t+s)}\right)\right] \\ &= e^{-\alpha t/2}\ e^{-\alpha(t+s)/2}\ e^{\alpha t} \\ &\qquad \text{because}\quad E[B(t_1)B(t_2)] = \min(t_1, t_2) \\ &= e^{-\alpha s/2} \end{aligned}$$

b. Suppose X is a Poisson random variable with mean λ. The λ itself is a random variable whose distribution is exponential with mean 1. Show that $P[X = n] = (0.5)^{n+1}$

Answer: Using the conditional expectation arguement, we have

$$\begin{aligned} P[X = n] &= \int_{-\infty}^{\infty} P[X = n \mid \lambda]\ f(\lambda)\ d\lambda \\ &= \int_{-\infty}^{\infty} \frac{\lambda^n}{n!}\ e^{-\lambda}\ f(\lambda)\ d\lambda \end{aligned}$$

$$
\begin{aligned}
& \quad \text{because, conditioned on } \lambda,\ X \text{ is Poisson} \\
= \ & \int_0^\infty \frac{\lambda^n}{n!}\, e^{-\lambda}\, e^{-\lambda}\, d\lambda \\
& \quad \text{because } \lambda \text{ is exponential with parameter } 1 \\
= \ & \int_0^\infty \frac{(2\lambda)^n}{2^n\, n!} e^{-2\lambda}\, \frac{1}{2} d(2\lambda) \\
= \ & \frac{1}{2^{n+1}\, n!} \int_0^\infty (2\lambda)^n e^{-2\lambda}\, d(2\lambda) \\
= \ & \frac{1}{2^{n+1}\, n!}\, \Gamma(n+1) \\
= \ & \frac{1}{2^{n+1}\, n!}\, n! \\
= \ & (0.5)^{n+1}
\end{aligned}
$$

E1 222 Stochastic Models and Applications
Assignment 3

1. Let X, Y be iid geometric random variables with parameter p. Let $Z = X - Y$ and $W = \min(X, Y)$. Find the joint mass function of Z, W. Show that Z, W are independent.

Answer: From their definitions, it is easy to see that W takes positive integer values while possible values for Z are both positive and negative integers (including zero). Also, when $Z > 0$ Y is smaller than X and when $Z < 0$, X is the smaller one.

Let $z > 0, w \geq 1$ be integers. Then

$$P[Z = z, W = w] = P[Y = w, X = w + z] = P[X = z + w]P[Y = w]$$

This gives us

$$f_{ZW}(z, w) = (1-p)^{z+w-1}p(1-p)^{w-1}p = (1-p)^{2w+z-2}p^2$$

Now consider the case where z, w are integers with $z < 0, w \geq 1$. Then

$$P[Z = z, W = w] = P[X = w, Y = w - z] = P[x = w]P[Y = w - z]$$

and hence

$$f_{ZW}(z, w) = (1-p)^{w-1}p(1-p)^{w-z-1}p = (1-p)^{2w-z-2}p^2$$

Finally when $z = 0$, we have

$$P[Z = z, W = w] = P[X = w, Y = w] = (1-p)^{w-1}p(1-p)^{w-1}p = (1-p)^{2w-2}p^2$$

Combining all these we now have

$$f_{ZW}(z, w) = (1-p)^{2w+|z|-2}\, p^2, w = 1, 2, \cdots, \ z \in \{\cdots, -2, -1, 0, 1, 2, \cdots\}$$

The marginal for Z is given by

$$f_Z(z) = \sum_{w=1}^{\infty} p^2(1-p)^{|z|}(1-p)^{2w-2} = p^2(1-p)^{|z|}\frac{1}{1-(1-p)^2} = \frac{p(1-p)^{|z|}}{2-p}$$

The marginal for W is given by

$$f_W(w) = \sum_{z=-\infty}^{\infty} p^2(1-p)^{2w-2}(1-p)^{|z|} = p^2(1-p)^{2w-2}\left(2\sum_{z=1}^{\infty}(1-p)^z + 1\right)$$

$$= p^2(1-p)^{2w-2}\left(\frac{2(1-p)}{p} + 1\right) = p(1-p)^{2w-2}(2-p)$$

From this, we can see that $f_{ZW}(z,w) = f_Z(z)f_W(w)$ and hence Z, W are independent.

2. Let X be a random variable having Gaussian density with mean zero and variance 1. Show that $Y = X^2$ has gamma density with parameters $\frac{1}{2}$ and $\frac{1}{2}$.
Now, let $X_1, \cdots, X_n$ be iid random variables having Gaussian density with mean zero and variance σ^2. Show that $Y = \frac{X_1^2+\cdots+X_n^2}{\sigma^2}$ has Gamma density with parameters $\frac{n}{2}$ and $\frac{1}{2}$. (This rv, Y, is said to have chi-squared distribution with n degrees of freedom).

Hint: The first part is solved in class. (See lecture 6). We also showed in class that if X, Y are independent Gamma rv with parameters α_1, λ and α_2, λ then $X + Y$ is gamma with parameters $(\alpha_1 + \alpha_2, \lambda)$.

3. Let X be uniform over $(0,1)$ and let Y be a discrete random variable taking non-negative integer values. Suppose X, Y are independent. let $Z = X + Y$. Show that Z is a continuous random variable.

Answer: To show that Z is a continuous rv, we need to show that F_Z is continuous everywhere and that it is differentiable at all but countably many points. (In an exam in this course, it is enough if you show continuity of F_Z).

Let $[z]$ denote the integer part of z (which is the largest integer less than or equal to z). We know that Z cannot take negative values and hence $F_Z(z) = 0$ for $z < 0$. Take $z \geq 0$.

$$\begin{aligned}
F_Z(z) &= P[X + Y \leq z] \\
&= P[Y \leq [z] - 1] + P[Y = [z], X \leq z - [z]], \quad \text{because } X \in (0,1) \\
&= \sum_{k \leq [z]-1} f_Y(k) + f_Y([z])\,(z - [z]) \\
&\quad \text{because } X, Y \text{ are independent and } X \sim U(0,1)
\end{aligned}$$

If z is not an integer, then, for sufficiently small δ, we would have $[z] = [z + \delta]$. Hence $F_Z(z + \delta) - F_Z(z)$ would be proportional to δ and hence F_Z would be continuous and differentiable at these points.

So, we only need to establish continuity at integer points. Also, df is always right continuous and hence we only need to show left continuity. Let $z = m - \delta$ where m is a positive integer and $\delta > 0$. Hence, $[z] = m - 1$. We need to show that $\lim_{\delta \downarrow 0} F_Z(z) = F_Z(m)$. Now, for this z,

$$\begin{aligned} F_Z(z) &= \sum_{k=0}^{m-2} f_Y(k) + f_Y(m-1)\,(1-\delta) \\ \text{and} & \\ F_Z(m) &= \sum_{k=0}^{m-1} f_Y(k) + f_Y(m)(m-m) \end{aligned}$$

which shows $\lim_{\delta \downarrow 0} F_Z(z) = F_Z(m)$.

4. Let X, Y, Z be iid continuous random variables. Show that $P[X < Y] = 0.5$ irrespective of what is the common density function of these random variables. Now calculate $P[X < Y < Z]$ and show that its value is same irrespective of what is the common density function of these random variables. Based on all this, can you guess what is the value of $P[X < Y, Z < Y]$. Explain.

Answer: We have already shown in class $P[X < Y] = 0.5$. Let f be the common density of the three random variables and let F be the corresponding distribution function. Then

$$\begin{aligned} P[X < Y < Z] &= \int_{-\infty}^{\infty} \int_{-\infty}^{z} \int_{-\infty}^{y} f_{XYZ}(x, y, z)\, dx\, dy\, dz \\ &= \int_{-\infty}^{\infty} \left(f(z) \int_{-\infty}^{z} \left(f(y) \int_{-\infty}^{y} f(x)\, dx \right) dy \right) dz \\ &= \int_{-\infty}^{\infty} \left(f(z) \int_{-\infty}^{z} f(y) F(y)\, dy \right) dz \\ &= \int_{-\infty}^{\infty} f(z) \frac{F(z)^2}{2}\, dz \\ &= \frac{1}{2}\,\frac{1}{3}(F(\infty)^3 - F(-\infty)^3) = \frac{1}{6} \end{aligned}$$

Given the above calculation, it is easy to see that $P[Z < Y < X]$ would also be the same. Thus all orderings of them would have equal probability. Hence, $P[X < Y, Z < Y] = P[X < Z < Y] + P[Z < X < Y] = \frac{2}{6}$

Comment If $X_1, X_2, \cdots, X_n$ are iid then any one specific ordering of them would have probability $\frac{1}{n!}$. This is because there are $n!$ orderings and all of them are equally likely.

5. Let $X_1, X_2, \cdots, X_n$ be random variables with mean zero and variance unity. Suppose the correlation coefficient of any pair of random variables, X_i and X_j, $i \neq j$, is ρ. Show that $\rho \geq \frac{-1}{n-1}$. Will this result remain true if $EX_i = \mu_i$ and $\mathrm{Var}(X_i) = \sigma_i^2$; but correlation coefficient between any pair of them is still ρ.

Answer: We have

$$\mathrm{Var}(X_1 + \cdots + X_n) = \sum_i \mathrm{Var}(X_i) + \sum_i \sum_{j \neq i} \mathrm{Cov}(X_i, X_j)$$

Hence, here we get

$$n + n(n-1)\rho \geq 0 \quad \Rightarrow \quad \rho \geq \frac{-1}{n-1}$$

If $X' = aX + b$ and $Y' = cY + d$ $(a, c > 0)$, then straight-forward algebra shows that $\rho_{X'Y'} = \rho_{XY}$. This shows that the answer to the second part of the question is yes because you can consider $X_i' = \frac{X_i - \mu}{\sigma}$ and so on.

6. Let X and Y be two discrete random variables with

$$P[X = x_1] = p_1, \quad P[X = x_2] = 1 - p_1;$$

$$P[Y = y_1] = p_2, \quad P[Y = y_2] = 1 - p_2.$$

Show that X and Y are independent if and only if they are uncorrelated. (Hint: Consider the special case where $x_1 = y_1 = 0$ and $x_2 = y_2 = 1$).

Answer: First consider the special case of X, Y taking values in $\{0, 1\}$ as suggested by the hint. We need to show (for this special case) that if the two random variables are uncorrelated then they are independent.

When X, Y are uncorrelated, we have $E[XY] = EX\ EY$. Given this we need to show that the joint mass function is product of the marginals.

When X, Y are binary, $E[XY] = P[X = 1, Y = 1]$, $EX = P[X = 1]$ and $EY = P[Y = 1]$. So, uncorrelatedness implies

$$P[X = 1, Y = 1] = P[X = 1]P[Y = 1] \quad \Rightarrow f_{XY}(1, 1) = f_X(1)f_Y(1)$$

Now we have:

$$P[X = 1, Y = 0] = P[X = 1] - P[X = 1, Y = 1] = P[X = 1] - P[X = 1]P[Y = 1]$$

$$= P[X = 1](1 - P[Y = 1]) = P[X = 1]P[Y = 0]$$

which shows $f_{XY}(1, 0) = f_X(1)f_Y(0)$. Similarly we can show the remaining two possibilities to complete the proof that X, Y are independent.

(The above is essentially the same as the fact that if two events A, B are independent then so are A, B^c, and A^c, B and so on).

Now consider the general case given in the problem. Define

$$X' = \frac{X - x_1}{x_2 - x_1}, \quad Y' = \frac{Y - y_1}{y_2 - y_1}$$

Since this is a linear (affine) transform, it is easily verified that X, Y being uncorrelated implies X', Y' are uncorrelated. This in turn implies that X', Y' are independent (because they take values in $\{0, 1\}$ which is the case proved above). By inverting the above transformations we can easily see that X is a function of X' and Y is a function of Y'. Hence X', Y' independent implies X, Y independent. This completes the solution of the problem.

7. An interval of length 1 is broken at a point uniformly distributed over $(0, 1)$. Let c be a fixed point in $(0, 1)$. Find the expected length of the subinterval that contains the point c. Show that this probability is maximized when $c = 0.5$.

Answer: Let X be uniform over $(0, 1)$. Let Y denote the length of the piece containing c. Y is a function of X. If $X > c$ (so that the rod is broken at a point to the right of c) then $Y = X$; otherwise (that is, if the rod is broken at a point to the left of c), $Y = (1 - X)$. Hence, we have $Y = XI_{[X \geq c]} + (1 - X)I_{[X<c]}$ where I_A denotes indicator of event A.

Let $g(X) = XI_{[X \geq c]}$. Then, $g(x) = x$ if $x \geq c$ and is zero otherwise. Hence $E[XI_{[X \geq c]}] = E[g(X)] = \int_0^1 g(x) f_X(x)\, dx = \int_c^1 x f_X(x)\, dx$. Also note that $f_X(x) = 1$ for $x \in (0, 1)$. Hence we get

$$EY = \int_c^1 x\, dx + \int_0^c (1 - x)\, dx = \frac{1}{2} + c - c^2$$

Now differentiating it with respect to c and equating to zero, you get $c = 0.5$. You can conclude this is a maximum by differentiating once more.