Probability & Statistics for EECS: Homework #07

Due on Dec 03, 2023 at 23:59

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

Problem 1

1. (a)

   When X is discrete, Y is discrete, we can have

$$P(Y=y|X=x)P(X=x) = P(X=x, Y=y) = P(X=x|Y=y)P(Y=y)$$
$$P(Y=y|X=x) = \frac{P(X=x|Y=y)P(Y=y)}{P(X=x)}$$

   Similarly, when X and Y are both continuous, we have

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y)$$
$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$

   When X is discrete, Y is continuous, suppose that there exist a positive number a

$$\lim_{a\to 0} P(Y \in (y-a, y+a)|X=x) = 2a \lim_{a\to 0} f_Y(y|X=x)$$
$$\lim_{a\to 0} \frac{P(X=x|Y \in (y-a,y+a)P(Y \in (y-a,y+a))}{P(X=x)} = \lim_{a\to 0} \frac{2aP(X=x|Y=y)f_Y(y)}{P(X=x)}$$

   By Bayes' rule

$$\lim_{a\to 0} P(Y \in (y-a,y+a)|X=x) = \lim_{a\to 0} \frac{P(X=x|Y \in (y-a,y+a)P(Y \in (y-a,y+a))}{P(X=x)}$$

   We can get

$$2a \lim_{a\to 0} f_Y(y|X=x) = \lim_{a\to 0} \frac{2aP(X=x|Y=y)f_Y(y)}{P(X=x)}$$

   Thus we conclude that

$$f_Y(y|X=x) = \frac{P(X=x|Y=y)f_Y(y)}{P(X=x)}$$

   When X is continuous, Y is discrete, similarly, we take a positive number a

$$P(Y=y|X=x) = \lim_{a\to 0} P(Y=y|X \in (x-a,x+a))$$

   By Bayes' rule

$$\begin{aligned}\lim_{a\to 0} P(Y=y|X \in (x-a,x+a)) &= \lim_{a\to 0} \frac{P(X \in (x-a,x+a)|Y=y)P(Y=y)}{P(X \in (x-a,x+a))} \\ &= \lim_{a\to 0} \frac{2af_X(x|Y=y)P(Y=y)}{2af_X(x)} \\ &= \frac{f_X(x|Y=y)P(Y=y)}{f_X(x)}\end{aligned}$$

   Thus we conclude that

$$P(Y=y|X=x) = \frac{f_X(x|Y=y)P(Y=y)}{f_X(x)}$$

Problem 1 continued on next page. . .

(b)

When X is discrete, Y is discrete

$$P(X = x) = \sum_{y} P(X = x, Y = y) = \sum_{y} P(X = x|Y = y)P(Y = y)$$

When X and Y are both continuous, we have

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)\, dy = \int_{-\infty}^{+\infty} f_{X|Y}(x|y) f_Y(y)\, dy$$

When X is discrete, Y is continuous, Using Bayes' Law

$$f_Y(y|X = x) = \frac{P(X = x|Y = y)f_Y(y)}{P(X = x)}$$

to get

$$\begin{aligned} P(X = x|Y = y)f_Y(y) &= f_Y(y|X = x)P(X = x) \\ \int_{-\infty}^{+\infty} P(X = x|Y = y)f_Y(y)dy &= \int_{-\infty}^{+\infty} f_Y(y|X = x)P(X = x) \\ &= P(X = x)\int_{-\infty}^{+\infty} f_Y(y|X = x)dy = P(X = x) \end{aligned}$$

Which means that

$$P(X = x) = \int_{-\infty}^{+\infty} P(X = x|Y = y)f_Y(y)dy$$

When X is continuous, Y is discrete, similarly, Using Bayes' Law

$$f_X(x|Y = y) = \frac{P(Y = y|X = x)f_X(x)}{P(Y = y)}$$

to get

$$\begin{aligned} P(Y = y|X = x)f_X(x) &= f_X(x|Y = y)P(Y = y) \\ \sum_y f_X(x|Y = y)P(Y = y) &= \sum_y P(Y = y|X = x)f_X(x) \\ &= f_X(x)\sum_y P(Y = y|X = x) = f_X(x) \end{aligned}$$

Which means that

$$f_X(x) = \sum_y f_X(x|Y = y)P(Y = y)$$

Problem 2

1. (a) If $x + y \neq n$, then

$$P(N = n, X = x, Y = y) = 0$$

, else

$$\begin{aligned} P(N = n, X = x, Y = y) &= P(N = n)P(X = x, Y = y|N = n) \\ &= P(N = n)P(X = x, Y = n - x|N = n) \\ &= \frac{\lambda^n e^{-\lambda}}{n!} \binom{n}{x} p^x (1-p)^{n-x} \end{aligned}$$

Since $X + Y = N$, so we can find that N, X, Y are not independent.

(b)

$$\begin{aligned} P(N = n, X = x) &= P(N = n, X = x, Y = n - x) \\ &= \frac{\lambda^n e^{-\lambda}}{n!} \binom{n}{x} p^x (1-p)^{n-x} \\ &\neq P(N = n)P(X = x) \end{aligned}$$

Since $N \sim Pois(\lambda)$ and $X \sim Pois(\lambda p)$, so X and N are not independent

(c)

$$\begin{aligned} P(X = x, Y = y) &= P(N = x + y, X = x) \\ &= \frac{\lambda^{x+ye^{-\lambda}}}{(x+y)!} \frac{\lambda^n e^{-\lambda}}{n!} \binom{x+y}{x} p^x (1-p)^y \\ &= \frac{(\lambda p)^x e^{-\lambda p}}{x!} \cdot \frac{(\lambda(1-p))^x e^{-\lambda(1-p)}}{y!} \\ &= P(X = x)P(Y = y) \end{aligned}$$

So X and N are independent

(d)

We can find from above that $X \sim Pois(\lambda p)$, $Y \sim Pois(\lambda(1-p))$, X and N are independent

$$\begin{aligned} Cov(N, X) &= Cov(X + Y, X) = Cov(X, X) + Cov(Y, X) = Var(X) = \lambda p \\ Corr(N, X) &= \frac{\lambda p}{Var(N)Var(X)} = \frac{\lambda p}{\sqrt{\lambda \lambda p}} = \sqrt{p} \end{aligned}$$

Problem 3

1. (a)

   If $t < x$, we will find that this situation is impossible. So

$$F_{T|X}(t|x) = 0$$

   if $t \geq x$,then

$$F_{T|X}(t|x) = F_{T|X}(t-x) = 1 - e^{-\lambda(t-x)}$$

   (b)

   Take the derivative of CDF, if $t < x$,

$$f_{T|X}(t|x) = 0$$

   if $t \geq x$,

$$f_{T|X}(t|x) = F'_{T|X}(t|x) = \lambda e^{-\lambda(t-x)}$$

   To verify it is a valid PDF, let's take the integral of $f_{T|X}(t|x)$ to get

$$\begin{aligned}\int_{-\infty}^{+\infty} f_{T|X}(t|x)dt &= \int_{-\infty}^{x} f_{T|X}(t|x)dt + \int_{x}^{+\infty} f_{T|X}(t|x)dt \\ &= 0 + \int_{x}^{+\infty} \lambda e^{-\lambda(t-x)}dt \\ &= \int_{0}^{+\infty} \lambda e^{-\lambda(t-x)}d(t-x) \\ &= -e^{-z}|_0^{+\infty} \\ &= -(0-1) \\ &= 1\end{aligned}$$

   (c)

   If $x \geq t$, similarly, we can find that

$$f_{X|T}(x|t) = 0$$

   else

$$f_{X|T}(x|t) = \frac{f_{T|X}(t|x)f_X(x)}{f_T(t)} = \frac{f_{T|X}(t|x)f_X(x)}{f_T(t)} = \frac{\lambda^2 e^{-\lambda t}}{f_T(t)}$$

   and since

$$\int_{-\infty}^{+\infty} f_{X|T}(x|t)dx = \int_{-\infty}^{t} f_{X|T}(x|t)dx + \int_{t}^{+\infty} f_{X|T}(x|t)dx = f_{X|T}(x|t)x|_0^t = tf_{X|T}(x|t) = 1$$

   So it is a valid PDF and we can find that if $x \leq t$,

$$f_{X|T}(x|t) = \frac{1}{t}$$

   (d)

   From (c), we got that

$$f_{X|T}(x|t) = \frac{\lambda^2 e^{-\lambda t}}{f_T(t)} = \frac{1}{t}$$

   when $x \leq t$ according to Baye's rule, thus, we can conclude that

$$f_T(t) = t\lambda^2 e^{-\lambda t}$$

Problem 4

1. (a)

   Suppose that $0 \leq m \leq 1$, then when $M \leq m$, it means that all of the three U_i are less than or equal to m, so we can find the CDF of M is

$$F_M(m) = m^3$$

   Then we can calculate that the PDF is

$$f_M(m) = 3m^2$$

   First we can find that $L \geq l$,$M \leq m$ means that all of the three U_i are between l and m, so we can find the CDF is

$$P(L \geq l, M \leq m) = (m-l)^3$$

   Since

$$P(M \leq m) = P(L \leq l, M \leq m) + P(L > l, M \leq m)$$

   Then the joint CDF of L, M is

$$\begin{aligned} P(L \leq l, M \leq m) &= P(M \leq m) - P(L > l, M \leq m) \\ &= m^3 - (m-l)^3 \end{aligned}$$

   the joint PDF of L, M is

$$f(l,m) = F'(l,m) = 6(m-l)$$

   (b)

   First we can find the CDF of L is

$$\begin{aligned} P(L \leq l) &= 1 - P(L > l) \\ &= 1 - P(U_1 > l, U_2 > l, U_3 > l) \\ &= 1 - (1-l)^3 \end{aligned}$$

   Then the PDF of L is

$$f_L(l) = 3(1-l)^2$$

   the conditional PDF of M given L is

$$\begin{aligned} f_{M|L}(m|l) &= \frac{f(l,m)}{f_L(l)} \\ &= \frac{2(m-l)}{(1-l)^2} \end{aligned}$$

Problem 5

1. (a)

$$\begin{aligned} f_{X,Y,Z}(x,y,z) &= f_{Y,Z|X}(y,z|x)f_X(x) \\ &= f_{Y|X}(y|x)f_{Z|X}(z|x)f_X(x) \\ &= \phi(y-x)\phi(z-x)\phi(x) \end{aligned}$$

(b) If there is not the condition, then we have no information of what distribution Y or Z is, thus we cann't conclude that Y and Z are also unconditionally independent.

(c)

$$f_{Y,Z}(y,z) = \int_{-\infty}^{+\infty} f_{Y,Z|X}(y,z|x)dx = \int_{-\infty}^{+\infty} \phi(y-x)\phi(z-x)\phi(x)dx$$

Problem 6

1. (a)

$$\begin{aligned} Cov(X,Y) &= E((X-\bar{x})(Y-\bar{y})) \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \\ &= r \end{aligned}$$

(b)

$E((X-\bar{x})(Y-\bar{y}))$ denotes the average signed area of the random rectangle formed by (X,Y) and $(\widetilde{X},\widetilde{Y})$

$$\begin{aligned} E((X-\bar{X})(Y-\bar{Y})) &= \frac{1}{n^2}(n\times 0+2\sum_{i<j}(x_i-x_j)(y_i-y_j)) \\ &= \frac{2S}{n^2} \end{aligned}$$

On the other hand, we can find that $E(XY) = E(\widetilde{X},\widetilde{Y})$,$E(\widetilde{X}Y) = E(X)E(\widetilde{Y}) = E(\widetilde{X})E(Y) = E(X)E(Y)$,then

$$\begin{aligned} E((X-\bar{x})(Y-\bar{y})) &= E(XY)+E(\widetilde{X}\widetilde{Y})-E(X\widetilde{Y})-E(\widetilde{X}Y) \\ &= E(XY)+E(XY)-E(X)E(Y)-E(X)E(Y) \\ &= 2(E(XY)-E(X)E(Y)) \\ &= 2Cov(X,Y) \end{aligned}$$

Thus, we can conclude that

$$Cov(X,Y) = \frac{S}{n^2}$$

(d)

(i)it does not matter that if we just swap the axis, since the area of the rectangles will not change.
(ii)This equation is equivalent to multiplying the length and width of a rectangle by respective factors, it results in an enlargement of the rectangle's area to the product of these two factors (like the formula for the area of a rectangle).
(iii)Adding or subtracting a constant is like shifting the rectangle a bit. Clearly, such a shift doesn't change the area of the rectangle.
(iv)Suppose that we can divide a rectangle with length l into two parts. One with length of x while the other with length l-x, which means that sum of the areas of these two smaller rectangles equals the area of the original whole rectangles.