Probability & Statistics for EECS: Homework #010

Due on Apr 23, 2023 at 23:59

Name: **Wang Penghao**
Student ID: 2021533138

Problem 1

(1) As for X discrete, Y discrete, we have that

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

(2) As for X continuous, Y discrete, we have that

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

Then we have that,

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

So we have that

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

(3) As for X discrete, Y continuous, as we have that

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

Then we have

$$P(X=x|Y=y)f_Y(y)=f_Y(y|X=x)P(X=x)$$

Then we integrate both sides with respect to y, we have that

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

Then we have that

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

So we get that

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

(4) As for X continuous, Y continuous, we have

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

So we get that

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

Problem 2

1. First, we let U be that the arrival time of the next Blissville company bus, then $U \sim Unif(0, 15)$ as the bus comes every 15 minutes, then we let $X \sim Expo(\frac{1}{15})$ is that be the arrival time of the next Blotchville company bus, then we have that

$$P(X < U) = \int_0^{15} P(X < U | U = u) \frac{1}{15} du$$

So we have that

$$P(X < U) = \frac{1}{15} \int_0^{15} (1 - e^{-\frac{u}{15}}) du = \frac{1}{e}$$

2. As for the wait time, that is the wait until the first bus comes. Denote that the wait time is W, then we have $W = min(X, U)$ Then as for the CDF, we firstly calculate the $P(W > t)$, then we have that

$$P(W > t) = P(X > t, U > t) = P(X > t)P(U > t) = e^{-\frac{t}{15}} (1 - \frac{t}{15})$$

So we get that the CDF of the waiting time is that

$$P(W \leq t) = 1 - P(W > t) = 1 - e^{-\frac{t}{15}} (1 - \frac{t}{15}),$$

where $t \in (0, 15)$ and CDF is 0 for $t \leq 0$ and 1 for $t \geq 15$.

Problem 3

(a) Firstly we denote that p is the probability that an egg hatch, and $q = 1 - p$. As we have that X is the number which hatch, and that Y is the number which do not hatch, we have that $N = X + Y$, then we have that N, X, Y are dependent, as N is the sum of two variables. Then we have that

$$
\begin{aligned}
P(N = n, X = x, Y = y) &= \sum_{n=0}^{\infty} P(X = x, Y = y | N = n) P(N = n) \\
&= P(X = x, Y = y | N = x + y) P(N = x + y) \\
&= P(X = x | N = x + y) P(N = x + y) \\
&= \binom{x+y}{x} p^x q^y \frac{e^{-\lambda} \lambda^{x+y}}{(x+y)!} \\
&= \frac{e^{-\lambda p} (\lambda p)^x}{x!} \frac{e^{-\lambda p} (\lambda p)^y}{y!}
\end{aligned}
$$

Where n, x, y are nonnegative integers and $n = x + y$. N, X, Y are not independent, but as X and Y are independent, we also get that $X \sim Pois(\lambda p)$, $Y \sim Pois(\lambda q)$

(b) As for the joint PMF of N, X, as from (a) we have that $X \sim Pois(\lambda p)$ and that $Y \sim Pois(\lambda q)$ Then we have

$$
P(N = n, X = x) = P(X = x, Y = n - x) = \frac{e^{-\lambda p} \lambda p^x}{x!} \frac{e^{-\lambda q} \lambda q^{n-x}}{(n-x)!},
$$

where $n \geq x$ and that X and N are dependent as $N \geq X$

(c) As for joint PMF of X, Y, from (a) we have that

$$
P(X = x, Y = y) = \frac{e^{-\lambda p} (\lambda p)^x}{x!} \frac{e^{-\lambda p} (\lambda p)^y}{y!}
$$

where x and y are nonnegative integers

(d) As for the relationship of X and N, from (a) we have that $X \sim Pois(\lambda p)$ and $Y \sim Pois(\lambda q)$, then we have that

$$
Cov(N, X) = Cov(X + Y, X) = Cov(X, X) + Cov(Y, X)
$$

As X and Y are independent, then

$$
Cov(N, X) = Var(X) = \lambda p
$$

We then have

$$
Corr(N, X) = \frac{Cov(X, Y)}{\sqrt{Var(X) Var(Y)}} = \frac{\lambda p}{\sqrt{\lambda \lambda p}} = \sqrt{p}
$$

Problem 4

Denote that the two measurements are X, Y, as they are 2 independent standard Normal random variables, so X, Y i.i.d. $\sim N(0,1)$ and denote that $M = max(X,Y)$, $L = min(X,Y)$
So $max(x,y) + min(x,y) = x + y$, and $max(x,y) - min(x,y) = |x - y|$, we then have that

$$E(M) + E(L) = E(M + L) = E(X + Y) = E(X) + E(Y) = 0$$

Then we have that as $X - Y = \sqrt{X}$ where $Z \sim N(0,1)$, and that $E(|X - Y|) = \sqrt{2}E|Z|$, we then have that By using 1D LOTUS, we have that

$$E|Z| = \int_{-\infty}^{\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 2\int_{0}^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \sqrt{\frac{2}{\pi}}$$

So we get that

$$E(M) - E(L) = E(M - L) = E(|X - Y|) = \frac{2}{\sqrt{\pi}}$$

So from $E(M) + E(L)$ and that $E(M) - E(L)$, we get that $E(M) = \frac{1}{\sqrt{\pi}}$, and that $E(L) = -\frac{1}{\sqrt{\pi}}$, then we also get that as $ML = XY$ and that $E(X) = E(Y) = 0$, we have that

$$Cov(M,L) = E(ML) - E(M)E(L) = E(XY) + \frac{1}{\pi} = 0 + \frac{1}{\pi} = \frac{1}{\pi}$$

.
Then as for $Var(M)$ and $Var(L)$, we have that as $E(X) = E(Y) = 0$, then

$$Var(X - Y) = E[(X - Y)^2] - (E[X - Y])^2 = E[(X - Y)^2] - (E(X) - E(Y))^2 = E[(X - Y)^2] = 2$$

As we also have that $X - Y = M - L$, then we have that

$$Var(X - Y) = E[(M - L)^2] = E[M^2] + E[L^2] - 2E[ML]$$

Then we have that

$$Var(X - Y) = E[M^2] + E[L^2] - 2E[x]E[Y] = E[M^2] + E[L^2]$$

As for the property of Variance, we have that $Var(M) = E(M^2) - EM^2$, the same for L. So we get

$$Var(X - Y) = Var(M) + EM^2 + Var(L) + EL^2 = Var(M) + Var(L) = \frac{2}{\pi}$$

Then as for the property of Normal distribution, we have $Var(M) = Var(L)$, we get that

$$Var(M) = Var(L) = (2 - \frac{2}{\pi})/2 = 1 - \frac{1}{\pi}$$

So we get the Corr is that

$$Corr(M,L) = \frac{Cov(M,L)}{\sqrt{Var(M)Var(L)}} = \frac{\frac{1}{\pi}}{1 - \frac{1}{\pi}} = \frac{1}{\pi - 1}$$

Problem 5

(a) As for the relationship between $Cov(X, Y)$ and sample variance, according to the definition of Cov, we have that

$$Cov(X, Y) = E([X - EX][Y - EY])$$

As $EX = \overline{x}$ and $EY = \overline{y}$, then we have that

$$Cov(X, Y) = E([X - \overline{x}][Y - \overline{y}]) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i = \overline{y}).$$

Which is the definition of sample covariance, so we get that

$$Cov(X, Y) = r.$$

(b) As for the total signed area of the rectangles, we have that the area is that

$$S = \sum_{i<j}(x_i - x_j)(y_i - y_j).$$

Then as for

$$E((X - \widetilde{X})(Y - \widetilde{Y})) = E(XY) + E(\widetilde{X}\widetilde{Y}) - E(X\widetilde{Y}) - E(Y\widetilde{X})$$

As according to the definition of the $\widetilde{X}$ and $\widetilde{Y}$, we have that they have the same distribution with X and Y. So, we have that as X and $\widetilde{Y}$, Y and $\widetilde{X}$ are independent,

$$E(X\widetilde{Y}) = E(Y\widetilde{X}) = E(X)E(Y)$$

and that

$$E(\widetilde{X}\widetilde{Y}) = E(XY)$$

So we have that

$$E((X - \widetilde{X})(Y - \widetilde{Y})) = 2E(XY) - 2E(X)E(Y) = 2Cov(X, Y)$$

1. As we have that

$$E((X - \widetilde{X})(Y - \widetilde{Y})) = \frac{0 * n + 2\sum_{i<j}(x_i - x_j)(y_i - y_j)}{n^2} = \frac{2A}{n^2}$$

2. As we have that

$$E((X - \widetilde{X})(Y - \widetilde{Y})) = 2Cov(X, Y),$$

we get that

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

(c) As for the following property, we have

(i) As W_1, W_2 are 2 r.v.s, then we as for exchange W_1, W_2, we consider that as exchange 2 axis in a coordinates, the space stay the same, so we have that

$$Cov(W_1, W_2) = Cov(W_2, W_1).$$

(ii) By mulitiply a constant to the axis, we can consider it as stretch the rectangle area on the corresponding axis, so that

$$Cov(a_1W_1, a_2W_2) = a_1a_2Cov(W_1, W_2)$$

Problem 5 continued on next page...

(iii) By adding a constant on a axis, this operation can be considered as do shift to the rectangle on the corresponding axis, so will not do any effect to the area. So we have that

$$Cov(W_1 + a_1, W_2, a_2) = Cov(W_1, W_2).$$

(iv) As for the $W_2 + W_3$, we can consider that it as the combination of 2 rectangles, so that whether we split it to 2 rectangles or in a single rectangle is the same area. So we have that

$$Cov(W_1, W_2 + W_3) = Cov(W_1, W_2) + Cov(W_1, W_3)$$