STAT2911 Assignment 2

Student ID: 480048691

May 21, 2020

1 Question 1

1.1 Find the marginal CDF of X

First, we want to find the marginal probability density function of X from the joint density, for all $x \in (0, \infty)$:

$$\begin{aligned} f_X &= \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_x^{\infty} e^{-(x+y)}dy = 2e^{-x}\int_x^{\infty} e^{-y}dy \\ &= 2e^{-x}\lim_{b\to\infty}\Big[-e^{-y}\Big]_x^b \\ &= 2e^{-2x}. \end{aligned}$$

So, we can derive the marginal CDF of X for $x \in (0, \infty)$:

$$F_X = \int_{-\infty}^{x} f_X(u)du = \int_0^x 2e^{-2u}du = 2\Big[-\frac{1}{2}e^{-2u}\Big]_0^x = 1 - e^{-2x}$$

1.2 Compute $\mathbf{P(Y < 1)}$ given $\mathbf{X < 1}$

By the definition of conditional probability,

$$P(Y < 1|X < 1) = \frac{P(Y < 1, X < 1)}{P(X < 1)}$$

Using the law of total probability, we have:

$$\begin{aligned} P(X < 1) &= P(X < 1, Y < 1) + P(X < 1, Y \geq 1) \\ &= \int_0^1\int_0^y f_{X,Y}(x,y)dxdy + \int_1^{\infty}\int_0^1 f_{X,Y}(x,y)dxdy \\ &= \int_0^1\int_0^y 2e^{-(x+y)}dxdy + \int_1^{\infty}\int_0^1 2e^{-(x+y)}dxdy \\ &= \int_0^1 2e^{-2y}dy + \int_1^{\infty}(-2e^{-1}+2)e^{-y}dy \\ &= (1-e^{-2}) + (-2e^{-2}+2e^{-1}) \\ &= 1 - 3e^{-2} + 2e^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} P(Y<1|X<1) &= \frac{P(Y<1, X<1)}{P(X<1)} \\ &= \frac{1-e^{-2}}{1-3e^{-2}+2e^{-1}} \\ &= \frac{e^2-1}{e^2+2e-3} && \approx 0.65024. \end{aligned}$$

1.3 Find the conditional density of Y given X

From above, we have found $f_X(x) = 2e^{-2x}$ for all $x \in (0, \infty)$.
Hence, for $x > 0$ and $y > x$:

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{2e^{-(x+y)}}{2e^{-2x}} = e^{x-y}.$$

1.4 Find the conditional expectation of Y given X

$$\begin{aligned} E(Y|X) &= \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_x^{\infty} y e^{x-y} dy \\ &= e^x \int_x^{\infty} y e^{-y} dy \\ &= e^x (x+1) e^{-x} \\ &= x+1. \end{aligned}$$

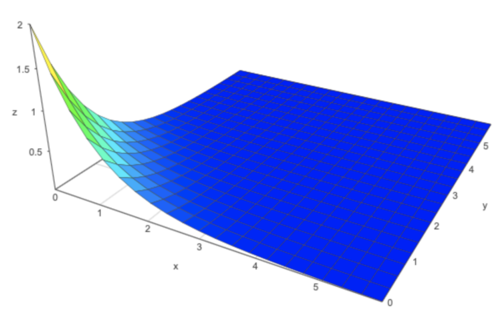
1.5 Compute P(Y < 1) given X = 1

Given the initial condition $x < y$ and $x = 1$, we know that y must be greater than 1, hence $P(Y < 1|X = 1) = 0$.

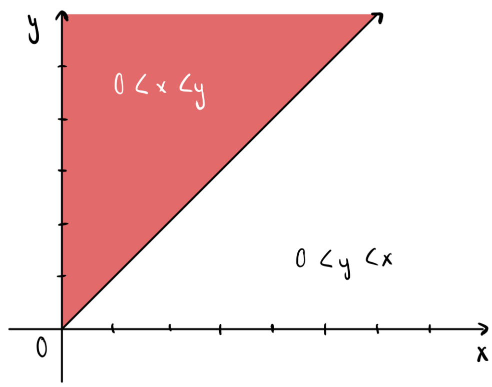
1.6 Joint CDF

Note that since $R_{XY} = \{(x,y)|0 < x < y\}$, $F_{XY}(x,y) = 0$, for $x \leq 0$ or $y \leq 0$.

Below is the joint density function of X and Y for $x, y > 0$:



However, note that the joint pmf is only defined over $0 < x < y$. That is, it is only defined over the shaded triangular region, when viewed in the xy-plane, as follows:



Given the constraints as above, we can see that:

• At $0 < x < y$ (the shaded region): The joint CDF of X and Y is the rectangular region (when viewed in the xy-plane) that gives the cumulative density up to $X \leq x, Y \leq y$. It does not depend on y until y is at least x. Hence,

$$\begin{aligned}
F_{XY} &= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{XY}(u,v) dv du \\
&= \int_{0}^{x} \int_{u}^{y} 2e^{-u-v} dv du \\
&= \int_{0}^{x} 2e^{-u} \Big(\int_{u}^{y} e^{-v} dv \Big) du \\
&= \int_{0}^{x} \big(-2e^{-u}e^{-y} + 2e^{-2u} \big) du \\
&= -2e^{-y} \int_{0}^{x} e^{-u} du + \int_{0}^{x} 2e^{-2u} du \\
&= 2e^{-y}e^{-x} - 2e^{-y} - e^{-2x} + 1 \\
&= 1 - 2e^{-y} - e^{-2x} + 2e^{-(x+y)}.
\end{aligned}$$

Indeed,

$$\begin{aligned}
\frac{\partial^2}{\partial x \partial y} F_{XY} &= \frac{\partial^2}{\partial x \partial y} \Big[1 - 2e^{-y} - e^{-2x} + 2e^{-(x+y)} \Big] \\
&= \frac{\partial}{\partial y} \Big[-2e^{-y}e^{-x} + 2e^{-2x} \Big] \\
&= 2e^{-(x+y)} = f_{XY}, \quad 0 < x < y.
\end{aligned}$$

• At $0 < y < x$ (the white region): The joint CDF of X and Y does not depend on x when x is greater than y. Hence,

$$
\begin{aligned}
F_{XY} &= \int_0^y \int_u^y f_{XY}(u,v)dvdu \\
&= \int_0^y \int_u^y 2e^{-u-v}dvdu \\
&= \int_0^y 2e^{-u}\bigg(\int_u^y e^{-v}dv \bigg)du \\
&= \int_0^y \big(-2e^{-u}e^{-y} + 2e^{-2u}\big)du \\
&= -2e^{-y}\int_0^y e^{-u}du + \int_0^y 2e^{-2u}du \\
&= 2e^{-2y} - 2e^{-y} - e^{-2y} + 1 \\
&= 1 - 2e^{-y} + e^{-2y}.
\end{aligned}
$$

Finally, the joint CDF of X and Y is:

$$
F_{XY} = \begin{cases} 0 & x \leq 0 \text{ or } y \leq 0, \\ 1 - 2e^{-y} - e^{-2x} + 2e^{-(x+y)} & 0 < x < y, \\ 1 - 2e^{-y} + e^{-2y} & 0 < y < x. \end{cases}
$$

2 Question 2

2.1 Expectation and Variance of the Sample mean

We have each $X_i \sim Uniform(0, \theta)$, hence $E[X_i] = \frac{\theta}{2}$ and $Var[X_i] = \frac{\theta^2}{12}$.
Expectation:

$$\begin{aligned}
E[\bar{X}] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E[X_i] \\
&= \frac{1}{n} nE[X_i] = \frac{\theta}{2}.
\end{aligned}$$

Variance:

$$\begin{aligned}
Var[\bar{X}] &= Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}Var\left[\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i] \\
&= \frac{1}{n^2} nVar[X_i] \\
&= \frac{1}{n}\frac{\theta^2}{12} = \frac{\theta^2}{12n}.
\end{aligned}$$

2.2 Expectation and Variance of the nth order statistic X_n

Given that $f(x) = \frac{1}{\theta}$ and $F(x) = \frac{x}{\theta}$

$$\begin{aligned}
f_n(x) &= \frac{n!}{(n-1)!(n-n)!} f(x)F(x)^{n-1}(1-F(x))^{n-n} \\
&= \frac{n!}{(n-1)!} f(x)F(x)^{n-1} \\
&= nf(x)F(x)^{n-1} \\
&= n\left(\frac{1}{\theta}\right)\left(\frac{x}{\theta}\right)^{n-1} \\
&= \frac{nx^n}{x\theta^n}.
\end{aligned}$$

Expectation

$$E[X_{(n)}] = \int_0^\theta x f_n(x)dx = \int_0^\theta x\frac{n}{x}\frac{x^n}{\theta^n}dx = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$$

Variance

$$E[X_{(n)}^2] = \int_0^\theta x^2 f_n(x)dx = \int_0^\theta x^2\frac{n}{x}\frac{x^n}{\theta^n}dx = \frac{n}{\theta^n}\frac{\theta^{n+2}}{n+2} = \frac{n\theta^2}{n+2}$$

Hence,

$$\begin{aligned}
Var[X_{(n)}] &= E[X_{(n)}^2] - E^2[X_{(n)}] \\
&= \frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1}\right)^2 \\
&= \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} \\
&= \frac{n\theta^2}{(n+1)^2(n+2)}.
\end{aligned}$$

2.3 Construct two unbiased estimators

We want to estimate θ using the sample mean and the n^{th} order statistic $X_{(n)}$.

Given the above results regarding the sample mean, we have:

$$\bar{X} = E[\bar{X}] = \tfrac{\theta}{2}, \text{ which makes } \tilde{\theta} = 2\bar{X}.$$

We can confirm this is an unbiased estimator since $E[\tilde{\theta}] = E[2\bar{X}] = 2E[\bar{X}] = \theta$.

Now, we need to find estimate $\hat{\theta}$ using the n^{th} order statistic such that $E[\hat{\theta}] = \theta$. Applying a linear function to $X_{(n)}$, we have the estimate $\hat{\theta} = cX_{(n)}$ for some constant c:

$$E[\hat{\theta}] = E[cX_{(n)}] = cE[X_{(n)}] = c\frac{n\theta}{n+1}$$

setting this equal to θ, we have:

$$c\frac{n\theta}{n+1} = \theta \quad \therefore \quad c = \frac{n+1}{n}$$

Hence,

$$\hat{\theta} = \frac{n+1}{n}X_{(n)}$$

2.4 Which estimator is better in the MSE sense?

Given that the found estimators are unbiased, we have

$$MSE(\tilde{\theta}) = Var[\tilde{\theta}] = Var[2\bar{X}] = 2Var[\bar{X}] = 2\frac{\theta^2}{12n} = \frac{\theta^2}{6n}.$$

$$\begin{aligned}
MSE(\hat{\theta}) &= Var(\hat{\theta}) \\
&= Var\Big[\frac{n+1}{n}X_{(n)}\Big] \\
&= \Big(\frac{n+1}{n}\Big)^2 Var[X_{(n)}] \\
&= \frac{(n+1)^2}{n^2}\frac{n\theta^2}{(n+1)^2(n+2)} \\
&= \frac{\theta^2}{n^2+2n}.
\end{aligned}$$

Since for $n \geq 5$:

$$\frac{\theta^2}{6n} > \frac{\theta^2}{n^2+2n}$$

so $\hat{\theta}$ is better in the MSE sense if the sample size is greater or equal to 5.