STAT 215A

Name: Quin Darcy | Due Date: 04/29/2022
Instructor: Dr. Cetin | Assignment: Test 02

1.A Let $X_1, X_2, \dots, X_n$ be n i.i.d. random variables, each with mean μ and standard deviation σ. Let $\overline{X} = \frac{X_1+X_2+\cdots+X_n}{n}$ represent the variable sample mean of $X_1, X_2, \dots, X_n$.

(a) Determine the expected value and variance of $\overline{X}$.

Solution. Since $X_1, \dots, X_n$ are independent (we may not need that X_k are independent and only need that E is a linear operator) and $E[X_k] = \mu$ for each $k = 1, \cdots n$, then

$$E\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} E[X_k] = n\mu.$$

Moreover, since for each $n \in \mathbb{N}$, $1/n \in \mathbb{R}$, then by properties of the expectation operator, we have that

$$E[\overline{X}] = E\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = \frac{1}{n}\sum_{k=1}^{n} E[X_k] = \frac{1}{n}(n\mu) = \mu.$$

To compute the variance of $\overline{X}$, we note that for each $k = 1, \dots, n$, $Var(X_k) = \sigma^2$. Moreover, since each $X_1, \dots, X_n$ are independent and therefore uncorrelated, then we have that

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) = n\sigma^2.$$

Since $1/n \in \mathbb{R}$ for every $n \in \mathbb{N}$, then by properties of the variance, it follows that

$$Var(\overline{X}) = Var\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \left(\frac{1}{n}\right)^2 \sum_{k=1}^{n} Var(X_k) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}.$$

■

(b) Let $Y_k = X_k - \overline{X}$ for each $k = 1, 2, \dots, n$. Determine the exected value and variance of Y_k for each $k = 1, 2, \dots, n$.

Solution. Since $Y_k = X_k - \overline{X}$ is a random variable and since the expectation operator is linear, then for any $k = 1, \dots, n$

$$E[Y_k] = E[X_k - \overline{X}] = E[X_k] - E[\overline{X}] = \mu - \mu = 0.$$

To compute the variance we first bring attention to several pieces. The first is that we were given that for all $k = 1, \dots, n$, the standard deviation of X_k is σ.

This implies that $\sqrt{E[X_k^2] - \mu^2} = \sigma$ and so $E[X_k^2] = \sigma^2 + \mu^2$ for every k. We also note that since $X_1, \dots, X_n$ are independent then for any $i, j \in \{1, \dots, n\}$ such that $i \neq j$ we have that X_i and X_j are independent and so

$$E[X_i \cdot X_j] = E[X_i]E[X_j] = \mu \cdot \mu = \mu^2.$$

The next thing to note is that for $k = 1, \dots, n$

$$\begin{aligned} Var(Y_k) &= E[(X_k - \overline{X})^2] \\ &= E[X_k^2 - 2X_k \cdot \overline{X} + \overline{X}^2] \\ &= E[X_k^2] - 2E[X_k \cdot \overline{X}] + E[\overline{X}^2]. \end{aligned}$$

By the first note made above, we already have a value for the first term $E[X_k^2]$, namely $\sigma^2 + \mu^2$. We now note that

$$\begin{aligned} X_k \cdot \overline{X} &= \frac{1}{n}(X_k X_1 + \cdots + X_k^2 + \cdots + X_k X_n) \\ &= \frac{1}{n}\left(X_k^2 + \sum_{j=1}^{k-1} X_k X_j + \sum_{j=k+1}^{n} X_k X_j \right). \end{aligned}$$

Since the expectation operator is linear, then from the above equality we obtain

$$\begin{aligned} E[X_k \cdot \overline{X}] &= E\left[\frac{1}{n}\left(X_k^2 + \sum_{j=1}^{k-1} X_k X_j + \sum_{j=k+1}^{n} X_k X_j \right) \right] \\ &= \frac{1}{n}\left(E[X_k^2] + \sum_{j=1}^{k-1} E[X_k]E[X_j] + \sum_{j=k+1}^{n} E[X_k]E[X_j] \right) \\ &= \frac{1}{n}\left(\sigma^2 + \mu^2 + \sum_{j=1}^{k-1} \mu^2 + \sum_{j=k+1}^{n} \mu^2 \right) \\ &= \frac{1}{n}(\sigma^2 + \mu^2 + (k-1)\mu^2 + (n-k)\mu^2) \\ &= \frac{n\mu^2 + \sigma^2}{n} = \mu^2 + \frac{\sigma^2}{n}. \end{aligned}$$

The last piece that we need to compute is $E[\overline{X}^2]$. Expanding the random variable we have

$$\overline{X}^2 = \frac{1}{n^2}(X_1 + \cdots + X_n)^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} X_i^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} X_i X_j \right).$$

From the above equation it follows that

$$\begin{aligned}
E[\overline{X}^2] &= E\left[\frac{1}{n^2}\left(\sum_{i=1}^{n} X_i^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} X_iX_j\right)\right] \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{n} E[X_i^2] + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} E[X_i]E[X_j]\right) \\
&= \frac{1}{n^2}\left(n(\sigma^2+\mu^2) + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} \mu^2\right) \\
&= \frac{1}{n^2}\left(n(\sigma^2+\mu^2) + 2\sum_{i=1}^{n} i\mu^2\right) \\
&= \frac{1}{n^2}\left(n(\sigma^2+\mu^2) + n(n+1)\mu^2\right) \\
&= \frac{(n+2)\mu^2+\sigma^2}{n}.
\end{aligned}$$

Putting all of these pieces together we get that

$$\begin{aligned}
Var(Y_k) &= E[(X_k^2 - \overline{X})^2] - E[Y_k]^2 \\
&= E[X_k^2] - 2E[X_k \cdot \overline{X}] + E[\overline{X}^2] \\
&= (\sigma^2+\mu^2) - 2\left(\mu^2 + \frac{\sigma^2}{n}\right) + \frac{(n+2)\mu^2+\sigma^2}{n} \\
&= \frac{2\mu^2 + (n-1)\sigma^2}{n}.
\end{aligned}$$

■

(c) Compute $Cov(Y_1, Y_2)$.

Solution. To compute the covariance, we will be using some of the pieces found in part (b). Letting $\mu_1 = E[Y_1]$ and $\mu_2 = E[Y_2]$, which are both zero by part (b), then we have that

$$\begin{aligned}
Cov(Y_1, Y_2) &= E[(Y_1-\mu_1)(Y_2-\mu_2)] \\
&= E[Y_1 \cdot Y_2] \\
&= E[(X_1-\overline{X})(X_2-\overline{X})] \\
&= E[X_1X_2 - X_1\overline{X} - X_2\overline{X} + \overline{X}^2] \\
&= E[X_1]E[X_2] - E[X_1\overline{X}] - E[X_2\overline{X}] + E[\overline{X}^2] \\
&= \mu^2 - 2\left(\mu^2 + \frac{\sigma^2}{n}\right) + \frac{(n+2)\mu^2+\sigma^2}{n} \\
&= \frac{2\mu^2 - \sigma^2}{n}.
\end{aligned}$$

■

2.A Let A be the upper half of the unit disk in $\mathbb{R}^2 : A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1 \text{ and } 0 \leq y \leq 1\}$. Assume that the joint pdf of two jointly continuous r.v. X and Y is given by

$$f_{X,Y}(x, y) = \begin{cases} C(1 + xy), & \text{if } (x, y) \in A \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine $C > 0$ so that $f_{X,Y}$ is a valid (joint) pdf.

Solution. To begin we note that for any $(x, y) \in A$ we have that $|xy| \leq 1$ and so $1 + xy \geq 0$. Moreover, if $C > 0$, then $C(1 + xy) \geq 0$ and so $f_{X,Y}$ satisfies the non-negative property of a jointly continuous pdf. Next, we need a value $C > 0$ such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1.$$

Since we are integrating over the upper half of the unit disk, we will convert to polar coordinates by letting $x = r \cos \theta$, $y = r \sin \theta$, and changing our bounds of integration to be $0 \leq r \leq 1$ and $0 \leq \theta \leq \pi$ since the integrals will evaluate to zero everywhere else. With this change of coordinates we have that

$$\begin{aligned} \int_0^{\pi} \int_0^1 C + Cr^2 \cos \theta \sin \theta \ \ dr d\theta &= C \int_0^{\pi} \left(\int_0^1 1 + r^2 \cos \theta \sin \theta \ \ dr \right) d\theta \\ &= C \int_0^{\pi} \left(1 + \frac{1}{3} \cos \theta \sin \theta \right) d\theta \\ &= C\pi - \frac{C}{3} \int_0^{\pi} u \ \ du \\ &= C\pi - \frac{C}{6} \cos^2 \theta \Big|_0^{\pi} \\ &= C\pi. \end{aligned}$$

Thus if $C\pi = 1$ then $C = 1/\pi$. ■

(b) Compute $P(X + Y > 1)$.

Solution. To begin, we want to define the region in which $x + y > 1$. It is clear that $x > 0$ since if $x \leq 0$, then this would mean $y > 1$ and such a point is not in A and thus has probability 0. Similarly, $x \leq 1$. Note that for any $0 < x \leq 1$, we need $y > 1 - x$. This means that the region of interest is the intersection of the plane $y > 1 - x$ and the upper half of the unit disc. To integrate over this region, note that for any $0 \leq \theta \leq \pi/2$, we have that

$$\sqrt{\cos^2 \theta + (1 - \cos \theta)^2} \leq r \leq 1.$$

Letting $f(\theta) = \sqrt{\cos^2\theta + (1-\cos\theta)^2}$, then we have that

$$\begin{aligned}
P(X+Y>1) &= \frac{1}{\pi}\int_0^{\pi/2}\int_{f(\theta)}^{1} 1 + r^2\cos\theta\sin\theta \;\; dr d\theta \\
&= \frac{1}{\pi}\int_0^{\pi/2} \big(1 - f(\theta) + \frac{1}{3}\cos\theta\sin\theta(1-(f(\theta))^3)\big) d\theta \\
&= \frac{1}{\pi}\Big[\int_0^{\pi/2} 1 - \sqrt{\cos^2\theta + (1-\cos\theta)^2} d\theta \\
&\qquad + \int_0^{\pi/2} \frac{1}{3}\cos\theta\sin\theta(1-(\cos^2\theta + (1-\cos\theta)^2)^{3/2}) d\theta\Big] \\
&\approx \frac{1}{\pi}(0.241253 + 0.0742742) \\
&\approx 0.100435.
\end{aligned}$$

■

(c) Determine the marginal pdf of X.

Solution. The marginal pdf of X is given by

$$f_X(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} 1 + xy \;\; dy = \frac{1}{\pi}\int_0^1 1 + xy \;\; dy = \frac{2+x}{4}.$$

■