APMA 1650 – Homework 7

1. Use the definition of expected value to show that if Y_1 and Y_2 are two continuous random variables with joint density function $f(y_1, y_2)$, then

$$E(Y_1Y_2) = E(Y_1)E(Y_2)$$

Note: Do not use Covariance for this!

$$\begin{aligned}
E(Y_1Y_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1y_2 f_{Y_1,Y_2}(y_1,y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1y_2 f_{Y_1}(y_1)f_{Y_2}(y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 && \text{by independence} \\
&= \int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2)\int_{-\infty}^{\infty} y_1 f_{Y_1}(y_1)\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&= \int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2)\left(\int_{-\infty}^{\infty} y_1 f_{Y_1}(y_1)\,\mathrm{d}y_1\right)\mathrm{d}y_2 \\
&= \left(\int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2)\,\mathrm{d}y_2\right)\left(\int_{-\infty}^{\infty} y_1 f_{Y_1}(y_1)\,\mathrm{d}y_1\right) \\
&= E(Y_1)E(Y_2)
\end{aligned}$$

2. Suppose X and Y are discrete random variables with joint probability mass function given by:

$$p(x,y) = \begin{cases} \frac{1}{3} & (x,y) = (-1,0), (0,1), (1,0) \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the covariance of X and Y?

We will do this using the Magic Covariance Formula:

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

First we find $\mathbb{E}(XY)$ using the pmf table above.

$$\mathbb{E}(XY) = (-1)(0)(1/3) + (0)(1)(1/3) + (1)(0)(1/3) = 0$$

Using the pmf table above, we can also find the marginal distributions of X and Y.

x	$p_X(x)$
-1	1/3
0	1/3
1	1/3

y	$p_Y(y)$
0	2/3
1	1/3

We can use these to compute the expected values of X and Y:

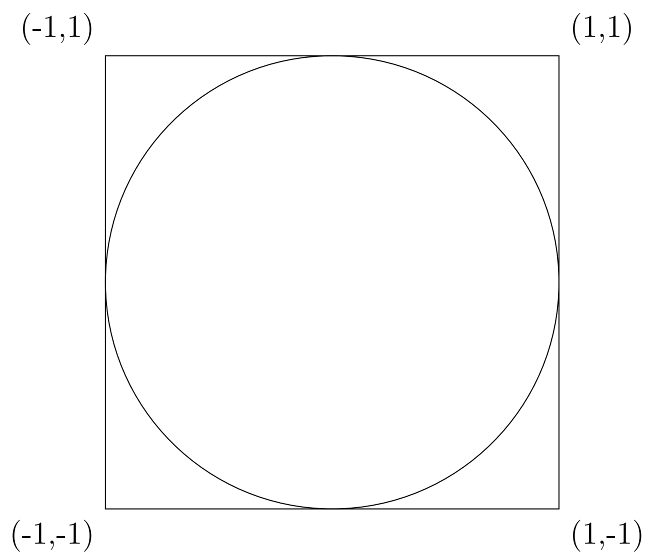
$$\mathbb{E}(X) = -1(1/3) + 0(1/3) + 1(1/3) = 0$$
$$\mathbb{E}(Y) = 0(2/3) + 1(1/3) = 1/3$$

By the Magic Covariance Formula, $Cov(X, Y) = 0 - (0)(1/3) = 0$.

(b) Are X and Y independent?

There are many ways to check that X and Y are not independent despite having a covariance of 0. One way is to note that Y is a function of X, i.e. the value of X completely determines Y. Alternatively, you can choose just about any x and y and show that $p(x, y) \neq p_X(x)p_Y(y)$.

3. Take the unit circle (circle of radius 1) and inscribe it in a square. See the diagram below.



You throw n darts uniformly at random at the square. Let Y be the number of darts which land in the circle. Form the estimator:

$$\hat{\theta} = \frac{4Y}{n}$$

(a) What is the expected value of $\hat{\theta}$? What is $\hat{\theta}$ the variance?

Y is the number of darts landing in the circle. The probability that a dart lands in the circle is the ratio of the circle area to the square area, which is $p = \pi/4$. Thus $Y \sim$Binomial$(n, \pi/4)$. Thus we have

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}\left(\frac{4Y}{n}\right) = \frac{4}{n}\mathbb{E}(Y) = \frac{4np}{n} = 4\frac{\pi}{4} = \pi$$

Thus this is an estimator for π. Alternatively, it is an estimator for the area of a circle.

For the variance use the formula for the variance of aX.

$$Var(\hat{\theta}) = Var\left(\frac{4Y}{n}\right) = \frac{16}{n^2}Var(X) = \frac{16}{n^2}\left(n\frac{\pi}{4}\right)\left(1-\frac{\pi}{4}\right) = \frac{4\pi - \pi^2}{n}$$

(b) In your own words, briefly explain how you can use this dart-throwing experiment to find an approximate value of π. Your answer should involve "Suppose you repeat that experiment many times..."

Suppose you repeat this experiment many times, and record the ratio of darts that land in the circle. If you then average the ratios of each experiment, part a tells us that this value should be "close" to π. As you increase n for each trial, and as you increase the number of trials, then the averaged value should get even closer to π yielding an approximate value of pi.

4. What do you know, another ACME widget factory question.

The new line of MiniWidgets you have launched have become so popular that your factory now has 10 machines devoted to producing MiniWidgets. The MiniWidgets produced by your machine have a mass which is normally distributed with a mean of 9 grams and a standard deviation of 0.2 grams.

(a) On a routine inspection, you take a sample of 9 MiniWidgets produced by a single MiniWidget machine. You find that the average mass of the MiniWidgets in your sample is 8.85 grams. Do you think the MiniWidget machine is defective? (Hint: compute the probability that the sample mean is less than or equal to 8.85 grams).

Let Y be the mass of one MiniWidget. Since the population has a normal distribution, the sample mean is normally distributed with mean 9 and standard deviation $0.2/\sqrt{9} = 0.2/3 = 1/15$. Thus we have:

$$\begin{aligned}\mathbb{P}(Y \leq 8.85) &= \mathbb{P}\left(Z \leq \frac{8.85 - 9}{1/15}\right) \\ &= \mathbb{P}(Z \leq -2.25) \\ &= 0.0122\end{aligned}$$

Thus it is highly likely that something is wrong with the machine.

(b) On another inspection, you find a machine which appears to be defective. You want to know the mean mass of widgets produced by that machine (population mean μ). To determine this, you take a sample of widgets from the machine and compute the sample mean $\bar{Y}$. Assume this machine produces widgets that are normally distributed, and that the standard deviation is the same as the other machines. How many widgets do you need to sample to be 95% confident that the sample mean is within 0.05 grams of the true population mean?

We can do this with the 68-95.99.7 rule. We want the true (population) mean to be within two (sample) standard deviations of the sample mean. Since we want two standard deviations of the sample mean to be 0.05, one standard deviation of the sample mean must be 0.025. Thus we solve:

$$\begin{aligned}\frac{\sigma}{\sqrt{n}} &= 0.025 \\ \sqrt{n} &= \frac{0.2}{0.025} = 8 \\ n &= 64\end{aligned}$$

Thus we should sample 64 MiniWidgets.

5. Let X be a binomial random variable with parameters n and p. Consider the following estimator for p:

$$\hat{p}_1 = \frac{X+1}{n+2}$$

(a) Find the bias of $\hat{p}_1$, which is defined as $E(\hat{p}_1) - p$.

By linearity of expectation,

$$\begin{aligned}\mathbb{E}(\hat{p}_1) &= \mathbb{E}\left(\frac{1}{n+2}X + \frac{1}{n+2}\right) \\ &= \frac{E(X)}{n+2} + \frac{1}{n+2} \\ &= \frac{np}{n+1} + \frac{1}{n+2} \\ &= \frac{np+1}{n+2}\end{aligned}$$

Thus we have:

$$\begin{aligned}Bias(\hat{p}_1) &= \mathbb{E}(\hat{p}_1) - p \\ &= \frac{np+1}{n+2} - p \\ &= \frac{np+1-p(n+2)}{n+2} \\ &= \frac{1-2p}{n+2}\end{aligned}$$

(b) Find the mean square error (MSE) of $\hat{p}_1$, which is bias squared plus the variance.

Since the MSE is bias squared plus variance, it suffices to find the variance. Using the formula for the variance of $aX + b$,

$$\begin{aligned} Var(\hat{p}_1) &= Var\left(\frac{1}{n+2}X + \frac{1}{n+2}\right) \\ &= \frac{1}{(n+2)^2}Var(X) \\ &= \frac{np(1-p)}{(n+2)^2} \end{aligned}$$

Adding the bias squared and the variance, we get the MSE:

$$\begin{aligned} MSE(\hat{p}_1) &= \left(\frac{1-2p}{n+2}\right)^2 + \frac{np(1-p)}{(n+2)^2} \\ &= \frac{(1-2p)^2 + np(1-p)}{(n+2)^2} \end{aligned}$$