Probability and Statistics—Exam 1
Thursday, November 18, 2021

Full Name: ____________________	Section 02D Mon., 1130 AM - 120 PM TA: Julio	Section 03D Mon., 130 PM - 320 PM TA: Li
Student ID Number: ____________________	Section 04D Mon., 330 PM - 520 PM TA: Li	Section 05D Wed., 1130 AM - 120 PM TA: Julio

- Write your full name and discussion section number on every page of this packet.
- **Show all work!** ... unless otherwise instructed. Partial credit can only be awarded for presented work. Full credit can only be awarded with presented work.
- You may use any calculator that does not have internet access (i.e. no smart phones, laptops, or tablets). Round approximate results to 4 decimal places.
- $\boxed{\text{Box}}$ your final answers.
- Uniformly distributed, each question is worth 10 points.
- You may use the back of this exam as scratch paper/additional space.
- Pages of formulas have been provided.

Full name: ______________________________ Section: __________

1. Bobcat Jensen is the new sensation among Twitch streamers. The following joint distribution shows how many Playstation games and how many Switch games they play during a randomly selected month. Compute the requested values for a randomly selected month.

		Switch		
		0	1	2
Playstation	0	0	0.11	0.06
	1	0.09	0.14	0.11
	2	0.09	0.12	0.12
	3	0.07	0.05	0.04

(a) probability that Bobcat Jensen played at least one Playstation game

Solution: Let X be the number of Playstation games and let Y be the number of Switch games that Bobcat Jensen played during a randomly selected month. Using the marginal probabilities,

		Switch			
		0	1	2	
Playstation	0	0	0.11	0.06	0.17
	1	0.09	0.14	0.11	0.34
	2	0.09	0.12	0.12	0.33
	3	0.07	0.05	0.04	0.16
		0.25	0.42	0.33	

$$\begin{aligned} P(X \geq 1) &= P(X=1)+P(X=2)+P(X=3) \\ &= 0.34+0.33+0.16 \\ &= 0.83 \end{aligned}$$

OR

$$\begin{aligned} P(X \geq 1) &= 1-P(X=0) \\ &= 1-0.17 \\ &= 0.83 \end{aligned}$$

(b) standard deviation of the number of Switch games played

Solution:

$$\begin{aligned} \mathrm{E}[Y] &= (0)(0.25)+(1)(0.42)+(2)(0.33) &&= 1.08 \\ \mathrm{E}[Y^2] &= (0)^2(0.25)+(1)^2(0.42)+(2)^2(0.33) &&= 1.74 \\ \mathrm{Var}(Y) &= \mathrm{E}[Y^2]-(\mathrm{E}[Y])^2 &&= 0.5736 \\ \mathrm{SD}(Y) &= \sqrt{\mathrm{Var}(X)} &&\approx 0.7574 \end{aligned}$$

(c) expected number of Playstation games played given that Bobcat Jensen played only one Switch game

Solution:

$$\begin{aligned} \mathrm{E}[X|Y=1] &= \sum_{i=1}^{4} x_i \cdot P(X|Y=1) \\ &= \frac{1}{P(Y=1)} \sum_{i=1}^{4} x_i \cdot P(X=x_i, Y=1) \\ &= \frac{(0)(0.11)+(1)(0.14)+(2)(0.12)+(3)(0.05)}{0.42} \\ &\approx 1.2619 \end{aligned}$$

Full name: ______________________________ Section: ________

2. A student in Math 32 may be tasked with completing a written homework assignment and a computer programming assignment in a week-long time frame. The units for times are in hours.[1] Let

   - $T_W \sim U(2,4)$ be the amount of time for a student to complete a written homework assignment, and
   - $T_C \sim N\left(\mu = \frac{1}{2}, \sigma^2 = \frac{1}{16}\right)$ be the amount of time for a student to complete a computer programming assignment

   Describe the distribution of time to complete both homework tasks by computing the mean and standard deviation of the sum $T_W + T_C$ assuming independence between T_W and T_C.

Solution: From the uniform distribution, the mean and variance of W are

$$\mu_W = \frac{2+4}{2} = 3$$

$$\sigma_W^2 = \frac{(4-2)^2}{12} = \frac{1}{3}$$

From the normal distribution, we are told that $\mu_C = \frac{1}{2}$ and $\sigma_C^2 = \frac{1}{16}$

From our study of linear operators,

$$\mathrm{E}[T_W + T_c] = \mu_W + \mu_C = 3 + \frac{1}{2} = 3.5 \text{ hours}$$

From the assumption of independence,

$$\mathrm{Var}(T_w + T_c) = \sigma_W^2 + \sigma_C^2 = \frac{1}{3} + \frac{1}{16} \approx 0.3958$$

and it follows that the requested standard deviation is

$$\sqrt{0.3958} \approx 0.6291 \text{ hours}$$

[1] Hint: there is only one input variable, time, so there is no need for double integrals.

Full name: ______________________________ Section: ________

3. Today, the student chefs are preparing the "Tortilla de Merced", a traditional Spanish egg dish with an infusion of spicy hot potato chips whose mascot is a cheetah.[2] Let us model the radius (in inches) of the dish with $R \sim \text{Exp}\left(\frac{1}{4}\right)$.

(a) Use Chebyshev's Inequality to compute how many dishes need to be measured so that their average radius is within 3 percent error with at least 95 percent probability.

Solution: We want to examine

$$P\left(\frac{|\bar{R}_n - \mu|}{\mu} < 0.03\right) = P\left(|\bar{R}_n - \mu| < 0.03\mu\right)$$

By Chebyshev's Inequality,

$$\begin{aligned} P\left(|\bar{R}_n - \mu| \geq 0.03\mu\right) &\leq \frac{\text{Var}(\bar{R}_n)}{a^2} &\leq 1 - 0.95 \\ P\left(|\bar{R}_n - \mu| \geq 0.03\mu\right) &\leq \frac{\sigma^2}{(0.03\mu)^2 n} &\leq 0.05 \end{aligned}$$

For an exponential distribution, $\sigma^2 = \mu^2$, and we can solve the inequality,

$$\begin{aligned} \frac{1}{(0.03)^2 n} &\leq 0.05 \\ \frac{1}{(0.03)^2(0.05)} &\leq n \end{aligned}$$

and arrive at $n \geq 22222.22$ dishes.

(b) Use the Central Limit Theorem to compute how many dishes need to be measured so that their average radius is within 3 percent error with at least 95 percent probability.

Solution: We want to examine

$$P\left(\frac{|\bar{R}_n - \mu|}{\mu} < 0.03\right) = P\left(-0.03\mu < \bar{R}_n - \mu < 0.03\mu\right)$$

[2] Hint: parts (a) and (b) will have different answers

Using the Central Limit Theorem, we scale by the standard error,

$$P\left(-\frac{0.03\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{R}_n - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{0.03\mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(-\frac{0.03\mu}{\frac{\sigma}{\sqrt{n}}} < Z_n < \frac{0.03\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

For an exponential distribution, $\sigma = \mu$, and we want the interval to encompass at least 95 percent probability. We can get a quantile with R code `qnorm(0.975)` and set

$$\begin{aligned} 1.96 &< \frac{0.03}{\frac{1}{\sqrt{n}}} \\ 1.96 &< 0.03\sqrt{n} \\ \frac{1.96}{0.03} &< \sqrt{n} \\ \left(\frac{1.96}{0.03}\right)^2 &< n \end{aligned}$$

and arrive at $n \geq 4268.444$ dishes.

Full name: ______________________________ Section: ________

4. Prove that if events X and Y are independent, then their correlation is zero. Your proof needs to include the definition of independence:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Solution: Beginning with the definition of independence, the joint expectation is

$$\begin{aligned}
\mathrm{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot P(X = x, Y = y)\, dy\, dx \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot P(X = x) \cdot P(Y = y)\, dy\, dx \\
&= \left(\int_{-\infty}^{\infty} x \cdot P(X = x)\, dx\right)\left(\int_{-\infty}^{\infty} y \cdot P(Y = y)\, dy\right) \\
&= \mathrm{E}[X] \cdot \mathrm{E}[Y]
\end{aligned}$$

Since $\mathrm{E}[XY] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$, the covariance is

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X] \cdot \mathrm{E}[Y] = 0$$

Since the covariance is zero, we conclude that the correlation,

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}} = \frac{0}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}} = 0$$