GROUND RULES:

- This exam contains 5 questions. Each question is worth 20 points. This exam is worth 100 points.
- This is a open-book and open-notes exam. You cannot use anything else other than the book or the notes; no internet; no discussion with anyone else except me, etc.
- Each question contains subparts. On each part, there is opportunity for partial credit, so show all of your work and explain all of your reasoning. **Translation:** No work/no explanation means no credit.
- If you use R to answer any part or to check your work, please include all code and output as attachments. Do not just write out the code you used.
- **My preference:** (1) Print out exam if you are attending the class online. (2) Print your name legibly at the top of this page. (3) Write your solution to each problem on its own page. You can use the back of the page if needed; you can append extra pages if needed. (4) Collect your pages when done and staple in upper left corner. Keep all pages/solutions in order.
- I prefer you write out your solutions "by hand." In the past, I have noticed students who type their solutions (in LaTeX or Word) often do not provide enough detail and/or do not explain their reasoning very well.
- Your solutions should be turned in to me no later than Tuesday, November 24 at 12:00 noon. **Note this change in time.** I would prefer your solutions be delivered to me in person. If this is not possible, let me know via email.

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

> *I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.*

1. Suppose Y_1 and Y_2 are continuous random variables with joint probability density function (pdf)

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \frac{20}{7}y_1^2y_2, & 0 \le y_1 \le 1,\ 1 \le y_2 \le y_1+1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Draw an excellent picture of the two-dimensional support

$$R = \{(y_1,y_2) : 0 \le y_1 \le 1,\ 1 \le y_2 \le y_1+1\}$$

and then show $f_{Y_1,Y_2}(y_1,y_2)$ integrates to 1 over R; i.e., $f_{Y_1,Y_2}(y_1,y_2)$ is a valid joint pdf.
(b) Are Y_1 and Y_2 independent? If so, prove it. If not, show why not.
(c) Calculate $P(Y_2 < 1.5)$.

2. The American Powerlifting Federation models the number of successful attempts in its weightlifting competitions. Two of the lifts in these competitions are the "bench press" and the "dead lift." Each competitor gets 3 attempts for each lift. Let

$$\begin{aligned} Y_1 &= \text{the number of successful bench press attempts (out of 3)} \\ Y_2 &= \text{the number of successful dead lift attempts (out of 3).} \end{aligned}$$

For competitors during the 2019 calendar year, the following table describes the joint probability mass function (pmf) of Y_1 and Y_2:

$p_{Y_1,Y_2}(y_1, y_2)$	$y_2 = 0$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$
$y_1 = 0$	0.01	0.02	0.01	0.01
$y_1 = 1$	0.01	0.06	0.12	0.06
$y_1 = 2$	0.02	0.08	0.30	0.10
$y_1 = 3$	0.01	0.04	0.05	0.10

(a) Draw an excellent picture of the two-dimensional support of Y_1 and Y_2.
(b) Find the marginal probability mass functions. Are Y_1 and Y_2 independent? Explain.
(c) Find the conditional pmf of the number of successful bench press attempts for competitors who are successful at exactly 1 dead lift attempt.
(d) Calculate the mean and variance of $Y_1 - 2Y_2$.

3. Suppose Y_1 and Y_2 are continuous random variables. The conditional probability density function (pdf) of Y_2, given $Y_1 = y_1$, is

$$f_{Y_2|Y_1}(y_2|y_1) = \begin{cases} y_1 e^{-y_1 y_2}, & y_2 > 0 \\ 0, & \text{otherwise,} \end{cases}$$

and the marginal distribution of Y_1 is uniform from 0 to 10; i.e., $Y_1 \sim \mathcal{U}(0, 10)$.

(a) Graph $f_{Y_2|Y_1}(y_2|y_1)$ for different values of $y_1 \in (0, 10)$. Label axes. Neatness counts.
(b) Derive the marginal pdf of Y_2. Make sure to note the support.
(c) Are Y_1 and Y_2 independent? Explain.
(d) Write $P(Y_1 - Y_2 > 0)$ as a double integral. You don't have to calculate it.

4. Suppose Z_1 and Z_2 are standard normal random variables; i.e., $Z_1 \sim \mathcal{N}(0,1)$ and $Z_2 \sim \mathcal{N}(0,1)$.

(a) If Z_1 and Z_2 are independent, calculate $P(Z_1^2 \leq 1, Z_2^2 \leq 1)$.
(b) If Z_1 and Z_2 are independent, calculate $P(Z_1^2 + Z_2^2 \leq 1)$.
(c) If Z_1 and Z_2 are independent ($\rho = 0$), determine the value of a that minimizes

$$V(aZ_1 + (1-a)Z_2).$$

Does you answer change when the correlation of Z_1 and Z_2 is $\rho \neq 0$?

5. Suppose X is a random variable with mean $E(X)$ and variance $V(X) = \sigma_X^2 > 0$. Suppose Y is a random variable with mean $E(Y)$ and variance $V(Y) = \sigma_Y^2 > 0$. Suppose we want to predict Y as a *linear* function of X. In other words, we are interested in functions of the form $Y = \beta_0 + \beta_1 X$, for fixed constants β_0 and β_1. Define

$$Q(\beta_0, \beta_1) = E\{[Y - (\beta_0 + \beta_1 X)]^2\}.$$

(a) Show that $Q(\beta_0, \beta_1)$ is minimized when

$$\beta_1 = \rho\left(\frac{\sigma_Y}{\sigma_X}\right)$$

and

$$\beta_0 = E(Y) - \beta_1 E(X),$$

where ρ is the correlation of X and Y. *Hint:* This is a minimization problem from multi-variable calculus.

(b) The prediction $Y = \beta_0 + \beta_1 X$, where β_0 and β_1 are given above, is the **best linear predictor** of Y. Calculate the equation of this line when the joint probability density function (pdf) of X and Y is

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{xy}{2}, & 0 < x < 2,\ 0 < y < x \\ 0, & \text{otherwise.} \end{cases}$$