Homework 6: Joint Distributions

In general, your goal in an explanation is to write enough that a student from class who has attended lecture, but not thought about the problem yet, could understand your approach, verify your reasoning, and believe your answer is correct. While we do not usually need to see arithmetic, you must include enough work that in principle one could rederive your answer with only a scientific calculator.

Unless a problem states otherwise, you should leave your answer in terms of factorials, combinations, etc., for instance 26^7 or $26!/7!$ or $26 \cdot \binom{26}{7}$ are all good forms for final answers.

Remember that you must "select pages" to identify which problem is on which page as part of your gradescope upload.

Submission: You must upload a **pdf** of your written solutions to Gradescope under "HW 6 [Written]". The use of is *required*.

Due Date: This assignment is due at **11 PM Tuesday August 8**.
You will submit the written problems as a PDF to Gradescope. Please put each numbered problem on its own page of the PDF (this will make selecting pages easier when you submit), and ensure that your PDFs are oriented correctly (e.g. not upside-down or sideways). The coding problem will also be submitted to Gradescope.

Note that there will be **no leeway** with the submission time. Any submission to Gradescope after 11:00:00 PM on the deadline day will be considered late. We recommend submitting a few minutes before the deadline in order to avoid technical issues with submissions.

Collaboration: Please read the full collaboration policy. If you work with others, you must still write up your solution independently and name all of your collaborators somewhere on your assignment.

Full submission instructions are in the course syllabus.

0. Extra Instructions :o

For calculations that require evaluating integrals (unless we indicate otherwise), you must

(a) Show the integral to evaluate (e.g., $\int_0^2 z \cdot 2dz$)

(b) Show an antiderivative and the values to evaluate at (e.g., $z^2|_0^2$)

(c) Plug in the values and simplify (e.g., $2^2 - 0^2 = 4$)

This is not a problem, so nothing needs to be submitted here.

1. I'm allergic to polling [6 points]

In lecture, we discussed an **i**dealized polling procedure and analysis to estimate the fraction p of a population that loves cats.

Specifically, we analyzed how we could choose the number n of people to sample in order to guarantee that 95% of the time, it will be the case that $\hat{p} = \bar{X} \in [p - .05, p + 0.05]$.

You probably know that many recent polls (that follow similar methodology) have been **way off**. Briefly discuss **two** reasons that real-world polling may not work as well as the idealized polling that we discussed.

Note: There can be many different answers. We expect most reasons will be 1-2 sentences.

2. Please I'm a Star [15 points]

An astronomer would like to measure the distance (in light years) from her observatory to a distant star. Each measurement is noisy, yielding only an estimate of the distance. Therefore, the astronomer plans to make a series of independent measurements and then use the average value of these measurements as her estimate of the actual distance. Suppose that each measurement has mean D (the true distance) and a variance of 7 light years. How many measurements does she need to make to be 96% confident that her estimate is accurate to within ± 0.3 light years? Your solution should use the CLT.

3. Take me to a Higher Level [25 points]

You watch dramas for H hours, where H is a random variable, equally likely to be 1, 2 or 3. The number of episodes E that you watch is random and depends on how long you watch for. (The episodes are of different lengths...). We are told that

$$\mathbb{P}(E = e \mid H = h) = \frac{1}{h}, \qquad \text{for } e = 1, \ldots, h \,.$$

(a) (5 points) Find the joint distribution of E and H.

(b) (5 points) Find the marginal distribution of E.

(c) (5 points) Find the conditional distribution of H given that $E = 1$ (that is, $\mathbb{P}(H = h \mid E = 1)$ for each possible h in 1,2,3). Use the definition of conditional probability and the results from previous parts.

(d) (10 points) Suppose that you watched 1 or 2 episodes. Find the expected number of hours you watched conditioned on this event, defined as follows:

$$\mathbb{E}\left[H \mid E = 1 \cup E = 2\right] = \sum_{h=1}^{3} h \cdot \mathbb{P}(H = h \mid E = 1 \cup E = 2)$$

4. Joint Density Ventures [12 points]

Let X and Y be continuous random variables with the following joint distribution.

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{2}x^2 y & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $\mathbb{P}(X > Y)$ [5 points].

(b) What is the marginal PDF of X? [3 points]

(c) What is the marginal PDF of Y? [3 points]

(d) Are X and Y independent? [1 point]

5. Don't Dimension It [10 points]

Let X, Y, Z be independent and uniformly distributed over (0,1) Compute

$$\mathbb{P}(X \ge YZ).$$

Your answer should be a number.

Hint: You just need to integrate over the relevant region of the joint density $f_{X,Y,Z}(x,y,z)$.

6. Law of Expected Employment [15 points]

12 students have decided to apply for a TA position next quarter. The number of courses that they can be a TA for is a Poisson random variable with mean 5. Suppose that each student independently chooses exactly one of the courses to apply for uniformly at random (and independently of the choices of the other applicants). Suppose also that each student has probability 0.2 of being acceptable as a TA to the professor teaching any course they apply for. Use the law of total expectation to compute the expected number of courses that have at least one applicant acceptable to the professor for that course. Your final answer should have only one summation symbol.