HOMEWORK 01

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

---

1. (**PSTAT 120A Review**) For a random variable X and constant a and b, show that

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

2. (**Modified from #5.36**) Let Y_1 and Y_2 denote the proportions of time (out of one workday) during which employees I and II, respectively, perform their assigned tasks. The joint relative frequency behavior of Y_1 and Y_2 is modeled by the density function

$$f_{Y_1,Y_2}(y_1, y_2) = (y_1 + y_2) \cdot \mathbb{1}_{\{0 \leq y_1 \leq 1,\ 0 \leq y_2 \leq 1\}}$$

   (a) Verify that this is a valid joint density function.

   (b) Find the marginal density functions for Y_1 and Y_2.

   (c) Are Y_1 and Y_2 independent? Why or why not? Be careful about your justification!

   (d) Find $\mathbb{P}(Y_1 \geq 1/2 \mid Y_2 \geq 1/2)$.

   (e) If employee II spends exactly 50% of the day working on assigned duties, find the probability that employee I spends more than 75% of the day working on similar duties.

3. (**Modified from #5.157**) A forester studying diseased pine trees models the number of diseased trees per acre, Y , as a Poisson random variable with mean λ. However, λ changes from area to area, and its random behavior is modeled by a gamma distribution. That is, for some integer α and a positive constant $\beta > 0$,

$$f(\lambda) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} \cdot \mathbb{1}_{\{\lambda \geq 0\}}$$

   Find the <u>unconditional</u> probability distribution of Y. Because α is assumed to be an integer, you should be able to recognize this distribution by name - include any/all relevant parameter(s)! **Hint:** When integrating/summing, try to multiply and divide by constants to get a density function inside the integral/sum. This will then avoid you having to perform any direct integration/summation!

4. Let N be a random variable whose support consists only of natural numbers, and let $\{Y_i\}_{i \geq 0}$ denote a sequence of identically distributed (but not necessarily independent) random variables with mean μ and variance σ^2. Furthermore, define

$$S_N := \sum_{i=1}^{N} Y_i$$

   Notice, crucially, that the sum on the RHS above contains a *random* number of terms.

   (a) Prove **Wald's Theorem**, which states that $\mathbb{E}[S_N] = \mathbb{E}[N] \cdot \mu$.

   (b) Derive a formula for Var(S_N). If you need to make an additional assumption, clearly state which assumption(s) need to be made.