Gov 2001: Problem Set 4

Spring 2026

Instructions:

- The problem set is due on **February 24, 11:59 PM Eastern Time**.
- Please upload a PDF of your solutions to Gradescope. Make sure to assign to each question all the pages with your work on that question.
- **Do not use AI assistants (ChatGPT, Claude, Copilot, etc.) on this problem set.** Work with each other instead. The struggle is where learning happens.
- Remember: 70% of your grade comes from in-class exams. Use problem sets to *learn*, not just to get answers.

Short Questions

1. Let X and Y be i.i.d. $\text{Unif}(0, 1)$. Let $S = X + Y$ and $D = X - Y$. Compute $\text{Cov}(S, D)$. Are S and D independent? You don't have to prove your answer, briefly explain in words.

2. Let X and Y be random variables with $\text{Corr}(X, Y) = \rho$. Let Z be independent of both X and Y. Let $A = X + Z$, $B = Y - Z$. Find $\text{Corr}(A, B)$ and express your answer in terms of ρ, $\text{Var}(X)$, $\text{Var}(Y)$, and $\text{Var}(Z)$.

Long Questions

3. In a neighborhood, suppose there is a latent "crime intensity" shock V during a given night (e.g., overall opportunity/guardianship conditions) that affects multiple types of incidents. Let W be the number of theft incidents driven by theft-specific factors, and let Z be the number of assault incidents driven by assault-specific factors. Let $X = V + W$ be the total observed counts of theft incidents and $Y = V + Z$ be the total observed counts of assault incidents.

As crime incidents are rare events, we assume V, W, Z are i.i.d. $\text{Pois}(\lambda)$.

(a) Find $\text{Cov}(X, Y)$.

(b) Are X and Y independent? Are they conditionally independent given $V = v$?

(c) Find the joint PMF of X, Y (as a sum).

Hint: Use the law of total probability over V:

$$\Pr(X = x, Y = y) = \sum_{v=0}^{\min(x,y)} \Pr(V = v) \Pr(W = x - v) \Pr(Z = y - v).$$

4. A researcher is trying to measure a group of voters' latent ideal points. Assume the true ideal point is $X \sim \mathcal{N}(\theta, 1)$. The researcher uses two different survey questionnaires, producing measurements Y and Z. Conditional on $X = x$, the measurement errors are independent and have Normal distributions:

$$Y = x + \varepsilon_1, \qquad Z = x + \varepsilon_2, \qquad \varepsilon_1 \sim \mathcal{N}(0, \sigma_1^2),\ \varepsilon_2 \sim \mathcal{N}(0, \sigma_2^2),$$

with $\varepsilon_1, \varepsilon_2$ independent of each other and of X.

(a) Find the joint PDF of X, Y, Z.

(b) By definition, Y and Z are conditionally independent given X. Discuss intuitively whether or not Y and Z are also unconditionally independent.

(c) Compute $\mathrm{Corr}(Y, Z)$. $\mathrm{Corr}(Y, Z)$ is often interpreted as the *consistency* of two survey measurements. Discuss how the measurement consistency changes when measurements become noisier.

5. In the last pset, we explored how Binomial-Poisson connection is useful for modeling the count of rare events (e.g. terrorist attack) in a cross-sectional setting. This question explores how Poisson distribution is related to Exponential distribution and how it helps model rare events in continuous time.

Consider you are studying vacancies on the U.S. Supreme Court. A vacancy occurs when a justice retires, dies, or resigns. As a first approximation, suppose the waiting time until the next vacancy is modeled as an Exponential random variable: $T \sim \mathrm{Exp}(\lambda)$, where T is measured in years and $\lambda > 0$ is constant over time.

(a) We've learned that Exponential distribution is memoryless: $\mathbb{P}(T > s + t \mid T > s) = \mathbb{P}(T > t)$. Explain in words how this property should be interpreted in the context of Supreme Court vacancies.

(b) A presidential term lasts 4 years. What is the probability that *no vacancy* occurs during the term?

(c) Suppose no vacancy has occurred during the first 2 years. What is the probability that *at least one vacancy* occurs during the remaining 2 years? Briefly explain why your answer is consistent with part (a).

Next we flip the question and ask: in a presidential term, how many vacancies could happen? Let $N(t)$ denote the number of vacancies occurring in the next t years, and suppose that after each vacancy, the waiting time until the next vacancy is again $\mathrm{Exp}(\lambda)$, independent of the past.

(d) Express the event $\{N(4) = 0\}$ in terms of the waiting time T and use this relationship to compute $\mathbb{P}(N(4) = 0)$.

(e) According to Poisson-Exponential connection, what's the distribution of $N(4)$? Use its PDF to compute $\mathbb{P}(N(4) = 0)$.

(f) Let $N_1 = N(2),\ N_2 = N(4) - N(2)$. Are N_1 and N_2 independent? What is the joint distribution of (N_1, N_2)?