Gov 2001: Midterm Exam

Spring 2025

March, 2025

Midterm Instructions:

- This midterm exam is due on **March 14, 11:59 pm** Eastern time. Please upload a PDF of your solutions to **Gradescope**. When submitting, please match your responses with the questions.
- We will accept hand-written solutions but we strongly advise graduate students to typeset your answers in LaTeX.
- This is a semi-closed book test. You are **NOT** allowed to: search internet / AI for solutions or communicate amongst each other.
- You are allowed to utilize class materials (slides, section slides, pset solutions).

1 Variance and Covariance (25pt)

1. Write down the definition of Covariance and interpret with no more than 2 sentences. (5pt)
2. Give a counter example when X and Y are dependent but with a covariance of zero. (5pt)
3. Prove the following (5pt each):
   - $\text{Cov}(X, X) = V(X)$
   - $\text{Cov}(X + Y, Z + W) = \text{Cov}(X, Z) + \text{Cov}(X, W) + \text{Cov}(Y, Z) + \text{Cov}(Y, W)$
   - $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$

2 Correlation (30pt)

1. Write down the definition of Correlation between two r.v.s X_1 and X_2 and show why it's always in between -1 and 1. (10pt)

   Hints: you will need to utilize Cauchy-Schwarz inequality: For any real-valued random variables X and Y),
   $$|E[AB]| \leq \sqrt{E[A^2]E[B^2]}.$$

2. Now, let's say you are studying the returns of two stocks, X_1 and X_2. Suppose the daily returns of these stocks are normally distributed with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ, meaning:

$$\text{Cov}(X_1, X_2) = \rho\sigma_1\sigma_2.$$

   (a) (10pt) Derive the expected return and variance of a portfolio consisting of an equal-weighted combination of these two stocks, i.e.,

$$S = \frac{X_1 + X_2}{2}.$$

   (b) Let's try to standardize the stock returns using:

$$Z_1 = \frac{X_1 - \mu_1}{\sigma_1}, \quad Z_2 = \frac{X_2 - \mu_2}{\sigma_2}.$$

      i. Compute the correlation coefficient between Z_1 and Z_2 (5pt).

      ii. Explain the significance of this correlation in the stock returns context with no more than 2 sentences. (5pt)

3 "Tea" Testing and Forecast (30pt)

In this problem, we're going to explore a real-world example of Fisher's "lady tasting tea" experiment from lecture: election forecasters – who have, for better or worse, become a big part of politics in the United States and elsewhere.

1. Suppose that Bob has correctly predicted six of the last eight election outcomes. What is the probability that someone randomly flipping a coin in each of the same elections would have experienced *at least* the same success as Bob? Compute your answer analytically (i.e. not by simulation). (10pt)

2. Forecasting has become so popular that riffraff are flooding the market. These "uniform amateurs" predict the vote share for each state in the U.S. presidential election by drawing a uniform random variable between 0 and 1, independently across states. You are deciding whether or not to hire a forecaster, Nate, to forecast each of the 50 state election winners in the 2024 presidential general election based on the performance of his 2020 election forecast, but you are worried that Nate might be one of these amateurs. When you ask him to justify his 2020 forecasts, he says "my highest predicted [Democratic] vote share was 0.8 which is very unlikely if I were a uniform amateur." Let's evaluate his claim.

   Suppose Nate is a uniform amateur and let X be the maximum of the 51 uniform vote share draws (include D.C.). Derive the CDF and PDF of X. Use these to calculate the probability of Nate's highest Democratic vote share being 0.8 or less if he were a uniform amateur. (10pt)

3. To be on the lookout for more uniform amateurs, it's helpful to know what highest vote share we should expect. To that end, calculate $E[X]$. (10pt)

4 Normal Distribution (15pt + bonus 10pt)

Let a standard normal r.v to be $Z \sim \mathcal{N}(0, 1)$.

1. Express the random variable $Y \sim \mathcal{N}(1, 2)$ as a simple function in terms of Z. Make sure to check that your Y has the correct mean and variance. (5pt)
2. Express the probability of $|Y| \leq 1$ as a function of Φ, the CDF of the standard Normal distribution. (10pt)

Bonus Q: Prove that if $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ and $X_1 \perp\!\!\!\perp X_2$,

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

(optional, 10pt)

Hint: You may use the moment-generating function (MGF) of a random variable X, denoted $M_X(t)$, defined as:

$$M_X(t) = E[e^{tX}].$$

The MGF uniquely determines the distribution. In other words, if X and Y are two r.v.s and for all values of t we have $M_X(t) = M_Y(t)$, then $F_X(k) = F_Y(k)$ for all values of k. To get the MGF of generalized $\mathcal{N}(\mu, \sigma^2)$, you can first derive the MGF of $\mathcal{N}(0, 1)$. You may also find this property helpful: If $S_n = \sum_{i=1}^{n} a_i X_i$, where the X_i are independent r.v.s and the a_i are constants, then $M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)$.