You have 150 minutes.

No notes, books, calculators, computers, etc. are allowed.

Show all of your work, in as organized a manner as possible. Incorrect answers with solid work often earn partial credit. Correct answers without explanatory work rarely earn full credit.

You may cite without proof any result discussed in class, the assigned textbook sections, or the assigned homework, unless a derivation is specifically requested.

Perform as much algebraic simplification as you can. Do not bother to do non-trivial arithmetic unless it is specifically requested. Mark your final answer clearly.

"log" denotes the natural logarithm — base $e \approx 2.718$.

Good luck. :)

**A1**. Assume a standard 52-card deck. A poker hand consists of 5 cards. A full house is a hand in which there are three cards of one rank and two cards of another rank — for example, three nines and two queens. What is the probability of being dealt a full house?

**A2**. Suppose that you have already been dealt four cards, which form two pairs. What is the probability that your fifth card will give you a full house?

**B1**. Recall that $X \sim \text{NegBin}(r, p)$ counts how many Bernoulli trials (independent with probability $p$) are needed to have $r$ successes. Explain how $X$ is a sum of independent geometric random variables.

**B2**. Find the moment generating function for $X$.

**C**. The disease seritits affects 1% of the population. There is a medical test for this disease, which is 95% accurate, meaning that 95% of people who have the disease test positive and 95% of people who don't have the disease test negative. I have just tested positive. What is the probability that I have the disease? (Do as much arithmetic simplification as possible, while keeping your answer exact.)

**D**. Suppose that $X$ and $Y$ are jointly distributed with density $f(x, y)$. They are not necessarily independent. Let $Z = X + Y$. Derive the density of $Z$ from first principles.

**E**. A biologist is studying a species of insect. Let $X_1, X_2, X_3, \ldots$ be the life spans, in days, of individuals from this species. She assumes that the $X_k$ are IID with unknown mean $\mu$ and variance $\sigma^2$. By picking a sample size $n$ and measuring $X_1, X_2, \ldots, X_n$, she can estimate $\mu$ and $\sigma^2$. But she wants to understand the uncertainty in her estimate of $\mu$.

Based on the central limit theorem, how large should $n$ be, so that she is 99% sure that her estimate of $\mu$ is within a given $\epsilon > 0$ of the true value? (My answer is in terms of $\epsilon$, the as-yet-unknown $\sigma^2$, and the CDF of the standard normal distribution.)

**F**. Explain how to use the inverse transform method to sample values of $X \sim \mathrm{Exp}(\lambda)$.

**G1**. Let $X$ be continuous with PDF $f$ and CDF $F$. Show that $f_{X|X>s}(x) = \dfrac{f(x)}{1-F(s)}$. (Hint: This is less confusing if you work with $F_{X|X>s}$.)

**G2**. Suppose now that $X$ represents the life span of a Joshco washing machine in years, and $f(x) = 0.7(0.2x-1)^6$ on $[0, 10]$. By the way, $E[X] = 5$. But what is the expected life span of a machine that has already been in service for one year?