Statistics 510, Homework 6

Brandon Legried

Fall 2021

Directions: Do all the problems. The total grade is 40 points, with all problems being worth 5 points each. You may work with others in the class or use any sources you like to study these problems, but when you go to write the solutions you should be able to write it using only the textbook and course notes. If you obtain facts, advice, or inspiration from any source outside the textbook or course notes, you must indicate in your solution what the source is. In all cases, you must make it clear to the reader what logic/work was needed to reach your solution. Solutions must be posted to Canvas – any legible version, handwritten or typed, is acceptable.

#1

Suppose that 26 percent of the students in a certain high school are freshmen, 24 percent are sophomores, 30 percent are juniors, and 20 percent are seniors.

(a) If 15 students are sampled with replacement at random from the school, what is the probability that at least eight will be *either* freshmen or sophomores?

(b) Following part a, let X_3 denote the number of juniors in the 15 sampled students and X_4 be the number of seniors in that sample. Compute $\mathbf{E}(X_3 - X_4)$ and $\mathrm{Var}(X_3 - X_4)$.

#2

Suppose we roll two (potentially unfair) six-sided dice A and B so that the outcomes from A and B are independent. Let S be the sum of the two rolls, supported on $\{2, 3, ..., 12\}$. Can the dice be constructed so that $\mathbf{P}(S = s) = 1/11$ for each $s \in \{2, 3, ..., 12\}$? Prove your answer.

#3

Let $X_1, ..., X_n$ be iid random variables, where $\overline{X}_n$ and S_n^2 are the sample mean and sample variance of the sequence.

(a) Show that

$$S_n^2 = \frac{1}{2n(n-1)} \sum_{i,j=1}^{n} (X_i - X_j)^2.$$

For the next two parts, assume that $\theta_1 = \mathbf{E}[X_i]$ and $\theta_j = \mathbf{E}\left[(X_i - \theta_1)^j\right]$ are finite. Note that θ_1 is the mean and θ_j are the second, third, and fourth centered moments.

(b) Show that $\text{Var}(S_n^2) = \frac{1}{n}\left(\theta_4 - \frac{n-3}{n-1}\theta_2^2\right)$.

(c) Find $Cov(\overline{X}_n, S_n^2)$ as a function of $\theta_1, \theta_2, \theta_3, \theta_4$. When does $Cov(\overline{X}_n, S_n^2)$ equal 0?

#4

Suppose that $X_1, ..., X_n$ are iid $Exp(1)$ random variables. Recall that $X_{(1)} = \min(X_1, ..., X_n)$ and $X_{(n)} = \max(X_1, ..., X_n)$. Determine the conditional pdf of $X_{(1)}$ given $X_{(n)} = y_n$.

#5

Let $X_1, ..., X_n$ be iid $Exp(\lambda)$ random variables and $Y_n = X_{(n)} - \log n = \max(X_1, ..., X_n) - \log n$.

(a) Show that Y_n converges in distribution to some random variable Y_∞; write the pdf of Y_∞.

(b) Find the mgf and expectation of Y_∞. For this, it will be useful to consider the positive constant

$$\gamma := \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log(n) \right).$$

You may use without proof that

$$\gamma = -\Gamma'(1),$$

where we recall the gamma function at each positive real number t is

$$\Gamma(t) = \int_{x=0}^{\infty} x^{t-1} e^{-x} \, dx.$$

(This γ is the *Euler-Mascheroni constant* and is approximately equal to 0.577; it appears in many places in analysis and probability.)

#6

(a) Let $p > 0$ be arbitrary and assume $\mathbf{E}|X_n|^p \to 0$ as $n \to \infty$. Show that X_n converges in probability to 0.

(b) Let $X_1, X_2, ...$ be uncorrelated random variables such that $\mathbf{E}[X_i] = \mu$ for all i and $\mathrm{Var}[X_i] \leq C < \infty$ for some C independent of i. Show that the sample average $\overline{X}_n$ converges in probability to μ.

#7

In this exercise, we consider improvements to the Chebyshev inequality and the law of large numbers for select examples.

(a) Let X be a random variable with moment generating function $M_X(s)$, defined in a neighborhood of 0. Show for every real number t that

$$\mathbf{P}(X \geq t) \leq \min_{s>0} e^{-st} M_X(s).$$

(b) Let $X \sim Bin(n, 0.5)$. Find a number $b > 1$ such that

$$b^n \mathbf{P}\left(\left|\frac{X}{n} - \frac{1}{2}\right| \geq \frac{1}{10}\right)$$

converges to a positive constant as $n \to \infty$.

#8

A physicist makes 25 independent observations of the specific gravity of a given body. Each measurement has a standard deviation of σ.

(a) Using the Chebyshev inequality, find a lower bound for the probability that the average of their measurements differs from the actual specific gravity by less than $\sigma/4$.

(b) Use the central limit theorem to approximate the probability in part (a).