STATS 116: Homework 6

Due: Friday, August 11, 2023 at 10:00 pm PDT on Gradescope

There are 7 problems on this assignment, each worth 8 points, although subparts within a problem may not be equally weighted. Credit will be assigned primarily based on reasoning and work, not the final answer. You do not need to simplify arithmetic expressions unless otherwise noted. While you may discuss the problems on this assignment other students, you must write up your own solutions. As per the syllabus, you may occasionally use the Internet or other public resources to clarify concepts with citation when this information is used as part of your own solution to a homework problem. However, you may not search for direct solutions to any problems assigned for homework or exams. For example, you can ask ChatGPT to clarify a particular concept from lecture that may be related to a problem, but you cannot feed it any part of a course assignment or a substantively similar version.

1. (BH 7.48) Athletes compete one at a time at the high jump. Let X_j be how high the jth jumper jumped, with X_1, $X_2, \dots,$ i.i.d. with a continuous distribution. We say that the jth jumper sets a record if X_j is greater than all of $X_{j-1}, \dots, X_1$. Find the variance of the number of records among the first n jumpers (as a sum). What happens to the variance as $n \to \infty$?

2. Suppose X_1 and X_2 are i.i.d. $\mathcal{N}(0, 1)$ random variables.

   (a) Find the PDF of X_1^2, and then use this to find the PDF of $Z = X_1^2 + X_2^2$. Show that Z has a Gamma distribution; specify the parameters.

   (b) Find the joint PDF of $Z_1 = X_1 + X_2$ and $Z_2 = X_1 - X_2$.

   (c) Are Z_1 and Z_2 independent? Are they uncorrelated? Explain.

3. Suppose $\lambda \sim \text{Expo}(\theta)$ for some $\theta > 0$ and that given λ, X has a Poisson distribution with rate parameter λ.

   (a) What is the marginal distribution of X?

   (b) Compute $\text{Cov}(X, \lambda)$. Hint: It may be helpful to use the law of iterated expectations, which will be covered in lecture on Monday, August 7. However, iterated expectations is really just LOTE written in a compact form, so the problem can be solved using LOTE.

4. (BH 7.71) In humans (and many other organisms), genes come in pairs. Consider a gene of interest, which comes in two types (alleles): type a and type A. The genotype of a person for that gene is the types of the two genes in the pair: AA, Aa, or aa (aA is equivalent to Aa). According to

the Hardy-Weinberg law, for a population in equilibrium the frequencies of AA, Aa, aa will be p^2, $2p(1-p)$, $(1-p)^2$ respectively, for some p with $0 < p < 1$. Suppose that the Hardy-Weinberg law holds, and that n people are drawn randomly from the population, independently. Let X_1;X_2;X_3 be the number of people in the sample with genotypes AA; Aa; aa; respectively.

(a) What is the joint PMF of X_1, X_2, X_3?

(b) What is the distribution of the number of people in the sample who have an A?

(c) What is the distribution of how many of the $2n$ genes among the people are A's?

(d) Now suppose that p is unknown, and must be estimated using the observed data X_1, X_2, X_3. The maximum likelihood estimator (MLE) of p is the value of p for which the observed data are as likely as possible. Find the MLE of p. Hint: recall $\log(x)$ is an increasing function.

(e) Now suppose that p is unknown, and that our observations can't distinguish between AA and Aa. So for each person in the sample, we just know whether or not that person is an aa (in genetics terms, AA and Aa have the same phenotype, and we only get to observe the phenotypes, not the genotypes). Find the MLE of p.

5. (BH 8.12) Let T be the ratio X/Y of two i.i.d. $\mathcal{N}(0,1)$ r.v..s X, Y. This is the *Cauchy* distribution and it has PDF

$$f_T(t) = \frac{1}{\pi(1+t^2)}$$

(a) Show that $1/T$ has the same distribution as T using calculus, after first finding the CDF of $1/T$ in terms of the CDF F_T of T.

(b) Argue that $1/T$ has the same distribution as T without using calculus.

6. (BH 8.18) Let X and Y be i.i.d. $\mathcal{N}(0,1)$ r.v.s, and (R,θ) be the polar coordinates for the point (X,Y) so $X = R\cos(\theta)$ and $Y = R\sin(\theta)$ with $R \geq 0$ and $\theta \in [0, 2\pi)$. Find the joint PDF of R^2 and θ. Also find the marginal distributions of R^2 and θ, giving their names (and parameters) if they are distributions we have studied before.

7. Emily and Devon each independently generate a $\text{Unif}(-1,1)$ number.

(a) Given that Emily's number is higher than Devon's, what is the expected value of Emily's number? Is this the same as the (unconditional) expected value of the maximum of Emily and Devon's numbers?

(b) What is the covariance between A, the maximum of Emily and Devon's numbers, and B, the minimum of Emily and Devon's numbers? Are A and B independent? Explain. Hint: Think along the lines of the identity $\max(a,b) + \min(a,b) = a + b$ for any numbers a, b.