Math 40 Probability and Statistical Inference
Homework 04

Winter 2021

Instructor: Yoonsang Lee (Yoonsang.Lee@dartmouth.edu)

Due Feb 8, 2021 11:59 pm (EDT)

Do the following exercises of the textbook. Show all steps to get your answers. Specify the problems you discussed with other students (including names).

5 points for each problem.

1. 3.1.5 Describe a bivariate random variable that has cdf $H(x, y) = \max(0, x) \times \max(0, y)$ for $x < 1$ and $y < 1$, and 1 elsewhere.

2. 3.1. 7 Let $G(x, y)$ and $H(x, y)$ be bivariate cdfs. Is $F(x, y) = \lambda G(x, y) + (1 - \lambda)H(x, y)$ a cdf $(0 \leq \lambda \leq 1)$? Answer this question by reducing to bivariate density.

3. 3.2.7 Following the derivation in Example 3.9, express $Pr(X > Y + 1)$ in terms of f and F.

4. 3.2.17 Find the distribution and the density function of the sum of **three** independent random variables uniformly distributed on (0,1).

5. 3.2.19 Random variables X and Y are independent and uniformly distributed on $(0, 1)$. Find the cdf of $Z = X/(X + Y)$. Hint: use a geometric approach.

6. (modified version of example 3.18) The probability that team A **loses** against team B is 0.6 and the probability that team C **wins** against team A is 0.7. What is the probability that team C **loses** against team B?

7. 3.3.8 The bivariate distribution is defined as the marginal pdf $X \sim \mathcal{E}(\lambda)$ and conditional pdf $Y|X = x \sim \mathcal{E}(\eta\lambda)$ for contants $\lambda, \eta > 0$. Find the joint density and $E(X|Y = y)$.

8. 3.3.13 Erica tosses a fair coin n times and Fred tosses $n + 1$ times. Prove that Fred gets more heads than Erica with probability 0.5 using the rule of repeated expectation.

9. 3.3.19 Random variables X and Y are uniformly distributed on the half disk bounded by $\sqrt{1 - x^2}$, where $|x| < 1$. Predict Y given $X = x$ and derive the conditional standard deviation of the prediction (this problem is NOT asking the standard deviation of Y given $X = x$).

10. 3.3.22 Prove that $Var(E(Y|X)) \leq Var(Y)$. What is your interpretation of this result? Discuss your interpretation in terms of data science (I will not take any questions about this problem).