

# Homework 10: 10/10 - 10/14

STA 335

Due 11:59pm Monday, October 24

Name: \_\_\_\_\_

**Instructions:** Write-up complete solutions to the following problems and submit answers on Gradescope. Your solutions should be neatly-written, show all work and computations, include figures or graphs where appropriate, and include some written explanation of your method or process (enough that I can understand your reasoning without having to guess or make assumptions). A rubric for homework problems appears on the final page of this assignment.

- Unless otherwise noted, problem numbers are taken from the 2nd edition of Blitzstein and Hwang's *Intro to Probability*.

## Monday 10/10

### Chapter 5

5, 13, 14

### Additional Problems

AP1. The *Pareto distribution* with parameter  $a > 0$  has PDF  $f(x) = a/x^{a+1}$  for  $x \geq 1$  (and 0 otherwise).

- Find the CDF of a Pareto r.v with parameter  $a$ .
- Pareto distributions are said to be *heavy-tailed*, which means they have relatively high probability of generating large values. For what values of  $a$  does a Pareto variable have a mean? A variance? Compute the mean and variance for those Pareto variables where it makes sense to do so.
- R does not have a formula for generating Pareto random variables (unlike `rbinom` for the binomial distribution). But R does have a function to generating Uniform random variable (`runif`). Explain how to use `runif` to generate 100 samples of a variable with the Pareto- $a$  distribution.

## Wednesday 10/12

### Chapter 5

22, 28, 32

### Additional Problems

AP2. We will prove later that if  $X, Y$  are independent with distribution  $N(\mu, \sigma^2)$ , then  $W = X + Y$  is  $N(2\mu, 2\sigma^2)$ . For now, we'll verify this empirically.

- Suppose  $\mu = 1, \sigma^2 = 4$ . Sample from the distribution of  $W$  1000 times by sampling from the distribution of  $X$  and from  $Y$  1000 times each, and then adding the values of  $X$  and of  $Y$  together. Compute the mean and variance of the sample of  $W$ .
- Create of histogram of the sample of  $W$ , and plot the density function for  $N(2, 8)$ . How do the histogram and density function compare?
- Create a QQ plot for  $W$  versus the theoretical quantiles for  $N(2, 8)$ . To do so, let  $a_1, a_2, \dots, a_{1000}$  be values of the sample of  $W$  in increasing order, and let  $F^{-1}$  be the quantile function for  $N(2, 8)$ . Plot the points  $(a_1, F^{-1}(1/1001)), (a_2, F^{-1}(2/1001)), \dots, (a_{1000}, F^{-1}(1000/1001))$ . Does the QQ-plot give evidence that  $W \sim N(2, 8)$ ?

## Friday 10/14

### Chapter 5

36, 37, 46

# Homework 10: 10/10 - 10/14

STA 335

Due 11:59pm Monday, October 24

Name:

## General Rubric

Points	Criteria
5	The solution is correct <b>and</b> well-written. The author leaves no doubt as to why the solution is valid.
4.5	The solution is well-written, and is correct except for some minor arithmetic or calculation mistake.
4	The solution is technically correct, but author has omitted some key justification for why the solution is valid. Alternatively, the solution is well-written, but is missing a small, but essential component.
3	The solution is well-written, but either overlooks a significant component of the problem or makes a significant mistake. Alternatively, in a multi-part problem, a majority of the solutions are correct and well-written, but one part is missing or is significantly incorrect
2	The solution is either correct but not adequately written, or it is adequately written but overlooks a significant component of the problem or makes a significant mistake.
1	The solution is rudimentary, but contains some relevant ideas. Alternatively, the solution briefly indicates the correct answer, but provides no further justification
0	Either the solution is missing entirely, or the author makes no non-trivial progress toward a solution (i.e. just writes the statement of the problem and/or restates given information)
<b>Notes:</b>	<p>For problems with multiple parts, the score represents a holistic review of the entire problem.</p> <p>Additionally, half-points may be used if the solution falls between two point values above.</p>