

General Notes

- You will submit a minimum of two files, the core files must conform to the following naming conventions (including capitalization and underscores). 123456789 is a placeholder, please replace these nine digits with your nine-digit Bruin ID. The files you must submit are:
 1. *123456789_stats102c_hw1.Rmd*: Your markdown file which generates the output file of your submission.
 2. *123456789_stats102c_hw1.html/pdf*: Your output file, either a PDF or an HTML file depending on the output you choose to generate.
 3. *Included image files*: If you answer your questions with images files, you must upload them to this portal as well, or your Rmd file will not knit.
 4. Please place all of your Rmd (and image) file(s) into a single folder named 123456789_stats102c_hw1 and compress the folder into 123456789_stats102c_hw1.zip.
 5. You will submit two files; one html/pdf file (123456789_stats102c_hw1.html/pdf) and one compressed file (123456789_stats102c_hw1.zip).

If you fail to submit any of the required core files you will receive **ZERO** points for the assignment. If you submit any files which do not conform to the specified naming convention, you will receive (at most) **half credit** for the assignment.

- **Your .Rmd file must knit.** If your .Rmd file does not knit you will receive (at most) half credit for the assignment.

The two most common reason files fail to knit are because of workspace/directory structure issues and missing include files. To remedy the first, ensure all of the file paths in your document are relative paths pointing at the current working directory. To remedy the second, simply make sure you upload any and all files you source or include in your .Rmd file.
- Your coding should adhere to the tidyverse style guide: <https://style.tidyverse.org/>.

NOTE: *Everything* you need to do this assignment is here, in your class notes, or was covered in discussion or lecture.

- Please **DO NOT** look for solutions online.
- Please **DO NOT** collaborate with anyone inside (or outside) of this class.
- Please work **INDEPENDENTLY** on this assignment.
- **EVERYTHING** you submit **MUST** be 100% your, original, work. Any student suspected of plagiarizing, in whole or in part, any portion of this assignment, will be **immediately** referred to the Dean of Student's office without warning.
- Only the base and stats packages are allowed to use for this assignment.

Problem 1: Suppose X is a geometric random variable with $p = 0.3$, determine the value of k such that $P(X \leq k) \simeq 0.8$

Problem 2: If two random variables, X and Y , are independent, show $\text{Var}(XY)$ in terms of the expected values and variances of X and Y .

Problem 3: At a call center in Los Angeles County, calls come in at an average rate of four calls per minute. Assume that the wait time from one call to the next follows the exponential distribution. Here we concern only with the rate of calling in and ignore the time spent on the phone. Also, assume that the waiting time between calls is independent. That is, a long delay between two calls doesn't make a shorter waiting period for the next call.

- (a) What is the average time between two successive calls?
- (b) What is the probability that exactly five calls occur within a minute?
- (c) What is the probability that less than five calls occur within a minute?
- (d) After a call is received, what is the probability that the next call occurs in less than 5 seconds?
- (e) Use **rexp** function in R to generate 1000 samples with the λ you derived from (d), and then draw a EDF plot with **ecdf** function. Finally, add confidence interval to your EDF plot.

Problem 4: Let X be a continuous random variable with cumulative density function $F(x)$. Let $u \sim \text{Unif}[0,1]$.

- (a) Let $Y = F^{-1}$. Show that $P(Y \leq x) = F(x)$.
- (b) Show that $F(x) \sim \text{Unif}[0,1]$.

Problem 5: Let $X \sim \text{two-parameter Exponential}(\lambda, \eta)$ with probability density function

$$f(x) = \lambda e^{-\lambda(x-\eta)}, x \geq \eta$$

, where λ and η are positive constant.

- (a) Find the cumulative density function $F(x)$
- (b) Write an algorithm using the CDF inverse method to generate random numbers from $\text{Exponential}(\lambda, \eta)$
- (c) Write R code to generate 10,000 random numbers and plot the histogram. You may set $\lambda =$ the last digit of your UID and $\eta =$ the second last digit. (If the digit is 0, please have it plus 1)
- (d) Compare the sample quantiles with the theoretical quantiles.

Problem 6: The Pareto distribution function has a probability density as shown below with the shape parameter α and the scale parameter β .

$$\frac{\alpha\beta^\alpha}{x^{\alpha+1}} \text{ for } x > \beta$$

- (a) Find the cumulative density function $F(x)$
- (b) Write an algorithm using the CDF inverse method to generate random numbers from $\text{Pareto}(\alpha = 5, \beta = 1)$
- (c) Write R code to generate 10,000 random numbers and draw a EDF plot.