HOMEWORK 2

Due February 15th, start of class

Your first homework covers Chapters 3.1, 3.2, 3.4.1, 3.4.2, 4.1, 4.2, and 4.3. You should be working on your homework throughout these first two weeks. If you can't solve some of the problems, please come to office hours. Email is fine only for very short questions.

THEORETICAL PORTION

The theoretical problems should be **neatly** numbered, written out, and solved. Do not turn in messy work.

1. You are on the show "Grab That Dough!" and are about to win some money! You are given a pot with that has 5 \$10 bills, 3 \$20 bills, 8 \$50 bills, and 2 \$100 bills. You are blind-folded and then allowed to pick two bills out of the pot (the bills are not put back in the pot after they are drawn). If you draw \$50 or less, you are then allowed to draw a third bill.

   Let X be the random variable denoting the sum of all three bills.

   (a) Write out the pdf and cdf of X.

   (b) What is the mean and variance of X?

   (c) It is going to cost you \$75 in gas money to drive to "Grab That Dough!" What is the probability that you earn that money back while on the show?

   (d) Do you think it is financially beneficial for you to go on "Grab That Dough!"?

2. Let X be a random variable such that $E(X) = 2, sd(X) = 2$, and let Y be a r.v. such that $E(Y) = 4, sd(Y) = 2$. X and Y are not linearly independent, and $corr(X, Y) = -0.3$.

   (a) What is $E(X + 3Y + 3)$?

   (b) What is $sd(X + 3Y + 3)$?

   (c) What is $E(3X - 2Y - 5)$?

   (d) What is $sd(3X - 2Y - 5)$?

3. A certain lab machine has 22 rings. With each use, these rings can fail, and an oil leak occurs. The probability of any ring failing during machine use is 8%. The rings are independent, and the failure of one ring does not impact the probability of failing for other rings. The machine gets serviced after each use, so any damaged rings are repaired after each use.

   (a) If 4 or more rings fail, the entire machine will shut down. What is the probability of the machine shutting down on any one use?

   (b) What is the probability that the machine runs successfully at least 5 times before shutting down?

4. (a) A traffic office wishes to monitor the number of vehicles crossing a certain bridge in the city in any given day. What family of distributions will they most likely be observing?

   (b) If it is estimate that an average of 30 cars cross the bridge per day, what is the probability that less than 10 cars will cross the bridge on any given day?

5. Answer the following questions regarding covariance and provide brief justification for your answer:

   (a) Can covariance between two random variables be less than -1?

   (b) If covariance between two random variables is negative, does their correlation have to be negative?

   (c) If $\text{Cov}(X, Y) = 0.3$, what is $\text{Cov}(100X, Y)$?

   (d) If $\text{Corr}(X, Y) = 0.1$, what is $\text{Corr}(100X, Y)$

   (e) What is $\text{Cov}(X, X)$?

(f) What is $\text{Corr}(X, X)$?

(g) What is $\text{Cov}(100X, 10X)$?

(h) What is $\text{Corr}(100X, 10X)$?

6. **APPM 5570 STUDENTS ONLY** Using the definition $EX = \sum_{x=0}^{\infty} xP(X = x)$, show that if $X \sim P(\lambda)$, then $EX = \lambda$.

COMPUTATIONAL PORTION

The computational portion of your homework should be neatly done and include all graphs, code, and comments, labeled and in order based on the problem you are addressing. Do *not* put graphs in at the end, stick code in random locations, or do anything else that will make this homework difficult to read and grade. **LABELS ARE YOUR FRIEND, USE THEM.** If you turn in something that is messy or out of order, it will be returned to you with a zero. All computations should be done using R, which can be downloaded for free at https://cran.r-project.org/.

1. Let X follow a geometric distribution, where the probability of success is 20%.

   (a) What is the probability that 7 failures occur before the first success?

   (b) What is the probability that 7 or more failures occur before the first success?

   (c) Plot this pdf for $X = 1, \ldots, 50$, using the X values on the x-axis, and $P(X = x)$ on the y-axis. To this same plot, add a red line showing the probabilities for $X = 1, \ldots, 50$ when the probability of success is 5%. What do you notice?

2. The function $f(x) = \frac{cx}{x^2+1}, 0 \leq x \leq 2$ is a pdf when the correct value for c is entered.

   (a) Use `integrate()` in R to solve for c in $f(x)$ above. Call this function $g(x)$.

   (b) Create a function called `calculate.Info()` such that for any given x the user enters, these components will be returned to the user for the function found in part (a):

      i. $g(x)$

      ii. $G(x) = \int_0^x g(u)du$, where $x \leq 2$.

      iii. **APPM 5570 students only:** The mean and standard deviation of the distribution.

   ALL STUDENTS: In one code file, email me `calculate.Info()` (and any other associated functions) by the time class starts on the due date. I will only enter the command:

```
calculate.Info(0.5)
```

   (or other numbers besides .05) into my console and will not do any de-bugging.

   A tip to help ensure your code will work on my machine:

   (a) Remove all objects from your R workspace. You can either do this by quitting and not saving anything or by typing `rm(list = ls())` into your console. You know all objects are removed if you type in `objects()` and get `character(0)` back.

   (b) Read in your function by typing `source('filepathname.r')` into your console, where 'filepathname.r' is an R file that includes only your function.

   (c) Type in `calculate.Info(0.5)` into your console. You should not get any syntax error messages.

3. The Weibull distribution: Generate random samples (5000 each) from four different Weibull distributions, with

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left\{-\left(\frac{x}{\beta}\right)^{\alpha}\right\}.$$

   - $\alpha = 1, \beta = 1$

- $\alpha = 3, \beta = 1$
- $\alpha = 1, \beta = 3$
- $\alpha = 3, \beta = 3$.

On one figure, plot the histograms (of the density, not the frequency) of all four samples, and add the theoretical density to each histogram for each beta distribution. Denote the theoretical distributions with red lines. Be sure to carefully label (using R) each plot.

4. The exponential distribution: generate random lightbulb failure times for 7500 lightbulbs from the exponential distribution with $\lambda = 0.75$.
   (a) What are the mean and standard deviation of the failure times? What should they be theoretically?
   (b) Using your sample, calculate the fraction of lightbulbs that failed in $(0, 2]$. What should this be theoretically?
   (c) Using your sample, for lightbulbs that lasted at least 2 years, calculate the fraction that failed in the next two years. Is the result what you expected?