M/STAT 501 Fall 2025: Homework 5

Due Wednesday, Oct 22 by 5:00pm in Gradescope

Directions: For this homework, there are three different types of problems: 1) *Graded for Accuracy Problems*, 2) *Graded for Completion Problems*, and 3) *Extra Practice Problems.* Please turn in your work for the *Graded for Accuracy Problems* and the *Graded for Completion Problems*, as well as any *Sources Used* by the due date via Gradescope. Show work and/or provide justification for credit—neatness counts! *The Extra Practice Problems* do not need to be turned in, and the solutions will be posted for you to check and compare your work.

You are encouraged to use R Markdown for your homework, but it is not required. However, **you must use R Markdown for any problems that require the use of R; include both your R code and the R output on the rendered pdf**.

Graded for Accuracy Problems (turn in)

1. For each of the following scenarios:
   (i) define an appropriate random variable (e.g., X = the number of heads in ten tosses of a fair coin),
   (ii) state whether the random variable is discrete or continuous,
   (iii) identify the name of an appropriate probability distribution to model the random variable, give the value(s) of its parameter(s), and explain your reasoning,
   (iv) find the required probability in R.

   a. The yield force of a steel reinforcing bar of a certain type is found to be normally distributed with a mean of 8500 pounds and a standard deviation of 80 pounds. If three such bars are to be used on a certain project, find the probability that all three will have yield forces in excess of 8600 pounds.

   b. Suppose that radioactive particles strike a certain target at an average rate of three particles per minute. Find the probability that 10 or more particles will strike the target in a particular two-minute period.

   c. The company that makes your external hard drive states that the average time to failure for their external hard drives is 5 years. You bought a 3 year warranty. What is the probability your external hard drive will fail before the warranty ends?

   d. Observations of songbirds indicate that shortly after dawn, they spend one to three hours feeding. A simple assumption is that the probability density function for feeding time is constant between 1 and 3. What is the probability a randomly chosen songbird spends less than 90 minutes feeding?

   e. A tire manufacturer designed a new tread pattern for its all-weather tires. Repeated tests were conducted on cars of approximately the same weight traveling at 60 miles per hour. The tests showed that the new tread pattern enables the cars to stop completely in an average distance of 125 feet with a standard deviation of 6.5 feet and that the stopping distances are approximately normally distributed. Suppose we take a random sample of 10 cars (of similar weight as the cars tested) equipped with these new tires and measure the stopping distance of each. Find the probability that at least eight of the ten sampled cars stop in less than 130 feet.

Graded for Completion Problems (turn in)

1. Explore the following references. If anything sparks your interest, explore further!
   - Skim through the Leemis & McQuestion article "Univariate Distribution Relationships"[1] (*The American Statistician*, 62(1), pp. 45–53), then explore their interactive distribution relationships at http://www.math.wm.edu/~leemis/chart/UDR/UDR.html.

   - Explore the "Common Families of Distributions in R" document and Rmd file (posted in Canvas → Modules → Chapter 3 Scaffolded Notes and R Examples) and try out the code using the R Markdown file of this document.

   - Explore and run the code in the `hw5-distribution-exploration.R` R script file (posted in Canvas → Modules → Homework Assignments).
     After working through these sources, post two discussion posts in the D2L Discussion topic "Probability Distribution Exploration":

   a. Create a new Canvas Discussions post in the "Probability Distribution Exploration" Discussion that answers the following three questions:
      - What is your favorite probability distribution and why?
      - Name one question that your probability distribution exploration brought up?
      - Name one other "fun fact" about a specific distributional family or relationships between families that you discovered besides the ones mentioned in class?

   b. Respond to another person's post commenting on any of their answers. Note that you must post first before you can see other posts.

2. Read the article "A Poisson Distribution and Poisson Process Explained"[2] by Will Koehrsen. Then answer the following questions.
   a. The article defines the rate parameter λ as follows (bold and italics in the original):

      > The most likely number of events in the interval for each curve is the **rate parameter**. This makes sense because the *rate parameter is the expected number of events in the interval* and therefore when it's an integer, *the rate parameter will be the number of events with the greatest probability.*

      The last claim—that the rate parameter is the number of events with the greatest probability—is stating that the mean (expected value) of a Poisson random variable is equal to the mode (value with the highest probability) when λ is an integer. Defining $X \sim Pois(\lambda)$ with pmf $f_X(x)$, use the expression

      $$\frac{f_X(x)}{f_X(x-1)} = \frac{e^{-\lambda}\frac{\lambda^x}{x!}}{e^{-\lambda}\frac{\lambda^{x-1}}{(x-1)!}} = \frac{\lambda}{x}$$

      to argue that the mean and the mode of a Poisson distribution will always differ by less than one.

   b. The article states that the waiting time between events in a Poisson process follows an exponential distribution with rate λ. That is, if T = time between two events in a Poisson process, then the pdf of T is

      $$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

      Prove this result. *Hint:* $P(T > t) = P(X = 0)$ where $X \sim Pois(\lambda t)$ (why?).

[1] https://www.stat.rice.edu/~dobelman/courses/texts/leemis.distributions.2008amstat.pdf
[2] https://productivityhub.org/2019/11/12/the-poisson-distribution-and-poisson-process-explained/

c. The following R function will simulate a Poisson process with rate `lambda` for a length of time `end.time`.

```
pois.process = function(lambda, end.time){
# Function to simulate a Poisson process
# with rate lambda over the interval [0, end.time].
# Returns:
# N = number of events
# T = event times
N = 0 # initialize event count
T = NULL # vector to store event times
t = rexp(1, rate = lambda) # waiting time to first event
if(t > end.time){
    return(list(N=0, T="NA"))
} else{
    while(t <= end.time){
        N = N+1
        T[N] = t
        t = t + rexp(1,lambda)
        }
    return(list(N=N, T=T))
}
}
```

For example, the following gives one simulation of a Poisson process with rate 2 over an interval $[0, 5]$:

```
set.seed(501)
my.process <- pois.process(2, 5)
my.process$N # Number of events
```

```
## [1] 16
```

```
my.process$T # Times of events
```

```
##  [1] 0.2181252 0.6576377 0.8891695 1.3282598 1.5026280 2.0534424 3.0430292
##  [8] 3.0602383 3.5291033 3.5433118 3.6358877 3.7088886 4.3349299 4.3741565
## [15] 4.6446238 4.9619194
```

Use this function to simulate the meteor random process described in the article 10,000 times—where we expect 5 meteors per hour on average and we watch for 1 hour. (Use hours as your unit of time.) Produce a histogram of your simulation like that shown in the article (figure with caption "Simulating 10,000 hours of meteor observations).

Challenge: (not required, but fun to try) Simulate this process five times. Plot the cumulative number of meteors over time for each run of the simulation on the same plot, where the x-axis is time, the y-axis is number of events, and each simulation is represented by one trajectory. *Hint*: Use the `plot` command with `type = "l"`, then add each additional simulation with the `lines` command.

Sources Used (turn in)

At the end of your assignment, list and cite **all** sources used when working on this assignment, including individuals with whom you discussed the problems, old homework assignments you may have found, discussion boards, websites, etc. If you did not use any sources besides our textbook or the course notes, please note this.

Extra Practice Problems (do not turn in, solutions posted in D2L)

1. The pdf of a normal distribution with mean μ and variance σ^2 is

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

for $-\infty < x < \infty$, where $-\infty < \mu < \infty$ and $\sigma^2 > 0$ are constants.

   a. Find the moment generating function (mgf) for a standard normal random variable (i.e., $\mu = 0$ and $\sigma^2 = 1$). *Hint*: Make the expression in the integral look like the pdf of a normal distribution. You will need to complete the square.

   b. Define $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0, 1)$. Find the mgf of the random variable X by taking advantage of the linear transformation $X = \sigma \times Z + \mu$.

2. Casella and Berger Exercises 3.12, 3.18, 3.20, 3.28, 3.46