M/STAT 501 Fall 2025: Homework 1

Due Wednesday, Aug 27 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 Wednesday, August 27 by 5:00pm 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.

Graded for Accuracy Problems (turn in)

1. We define the symmetric difference of two sets A and B as

$$S(A, B) = (A \cap B^C) \cup (A^C \cap B).$$

   Show that

$$P(S(A, B)) = P(A) + P(B) - 2P(A \cap B).$$

2. Prove Bonferroni's Inequality: $P(A \cap B) \geq P(A) + P(B) - 1$. Then verify Bonferroni's Inequality using Venn diagrams. (Hint: For the Venn diagram, consider the rearranged identity $1 \geq P(A) + P(B) - P(A \cap B)$.)

3. A manufacturing company has two retail outlets. It is known that 25% of all potential customers buy products from outlet 1 only, 35% buy from outlet 2 only, and 8% buy from both 1 and 2. Let A denote the event that a potential customer, randomly chosen, buys from outlet 1, and let B denote the event that the customer buys from outlet 2. Suppose a potential customer is chosen at random. For each of the following events,

   - represent the event symbolically in terms of A and B,
   - draw a Venn diagram with the event shaded, and
   - find its probability.

   a. The randomly chosen customer buys from outlet 1.
   b. The randomly chosen customer does not buy from outlet 2.
   c. The randomly chosen customer does not buy from outlet 1 *or* does not buy from outlet 2.
   d. The randomly chosen customer does not buy from outlet 1 *and* does not buy from outlet 2.

4. Let A and B be events with probabilities $P(A) = \frac{3}{4}$ and $P(B) = \frac{1}{3}$.

   a. Show that $\frac{1}{12} \leq P(A \cap B) \leq \frac{1}{3}$.
   b. Find corresponding bounds for $P(A \cup B)$.

Graded for Completion Problems (turn in)

We will be using R and RStudio throughout the semester. For this first assignment, based on your familiarity with the software, you may want to start with the optional tutorial described in 1). If you are comfortable using R and don't need a review, please start with the activity described in 2). Please don't spend hours fretting over code that just won't work–if you get stuck, post your code and questions in our Canvas Discussion

board. This is meant to be introductory activity, and your solutions for this portion will be graded only on completion, not accuracy.

1. (*Optional*) **Beginning R or a little rusty?** Complete the "Getting Started with R" introduction (available under Assignments -> Homework -> Homework 1 in Canvas). This introduction will have you install R and RStudio, as well as practice writing R code. You do not need to turn in a copy of your R code for this tutorial.

2. **Ready to dive right in?** Use simulation in R to estimate the probability found for book problem 1.22(b). How does this estimate compare to the exact probability shown? (We will derive this probability in the next homework.) Please turn in a printed copy of relevant code and output that have been neatly incorporated into a single document. Examples of functions/code are provided below. (*Optional challenge:* Use simulation to estimate the probability found for book problem 1.22(a).)

```r
# Calculate (40 choose 10)
choose(40,10)

# Randomly select 6 letters from the alphabet without replacement.
sample(letters, size=6, replace=FALSE)

# Randomly select 6 letters from the alphabet with replacement.
sample(letters, size=6, replace=TRUE)


## Use simulation to estimate the probability of getting an 'A'
## in a sample of size 6 for the without replacement sampling strategy.

# Step 1: Write a function to simulate a random sample of
# n letters and record whether "A" is observed
sample.function <- function(n, replacement) {
  num.A <- sum(sample(LETTERS, size = n, replace = replacement) == "A")
  if (num.A >= 1) {
indicator <- 1
  } else {
indicator <- 0
  }
}

# Step 2: Simulate 10,000 random samples each of size n = 6 letters,
# without replacement
sim.wo <- replicate(10000, sample.function(6, FALSE))

# Step 3: Estimate probability by finding the proportion of the 10,000
# samples that had an "A"
est.prob.wo < mean(sim.wo)
```

3. What unresolved questions, concerns and/or issues did you encounter while working in RStudio for this homework? Post and respond to these questions on Canvas.

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. 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. Give a sample space for each of the following situations:
   a. The number of calls that arrive at a customer service hotline in an hour is counted.
   b. A customer service representative starts work at 9:00 am and works until 5:00 pm. The elapsed time is measured until the representative receives the first call of his shift.
2. For any three events A, B and C in S, prove that
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$
3. Casella and Berger Exercise 1.4.
4. Casella and Berger Exercise 1.11 parts (a) and (c).