# Math 151 - Probability Theory - Homework 4

your name here

Due: Friday, September 18, 2020, midnight PDT

## Important Note:

You should work to turn in assignments that are clear, communicative, and concise. Part of what you need to do is not print pages and pages of output. Additionally, you should remove these exact sentences and the information about HW scoring below.

Click on the *Knit to PDF* icon at the top of R Studio to run the R code and create a PDF document simultaneously. [PDF will only work if either (1) you are using R on the network, or (2) you have LaTeX installed on your computer. Lightweight LaTeX installation here: https://yihui.name/tinytex/]

> Either use the college's RStudio server (https://rstudio.pomona.edu/) or install R and R Studio on to your personal computer. See: https://research.pomona.edu/johardin/math151f20/ for resources.

#### Assignment

#### 1: PodQ

Describe one thing you learned from someone in your pod this week (it could be: content, logistical help, background material, R information, etc.) 1-3 sentences.

#### 2: 2.1.12

For any three events $A, B$, and $D$, such that $P(D) > 0$, prove that

$$P(A \cup B|D) = P(A|D) + P(B|D) - P(AB|D)$$

#### 3: 2.2.2

Assuming that $A$ and $B$ are independent events, prove that the events $A^c$ and $B^c$ are also independent.

#### 4: 2.2.10

The probability that any child in a certain family will have blue eyes is 1/4, and this feature is inherited independently by different children in the family. If there are five children in the family and it is known that at least one of these children has blue eyes, what is the probability that at least three of the children have blue eyes?

#### 5: 2.2.11

Consider the family with five children described in Exercise 10 (from section 2.2 in DeGroot).

a. If it is known that the youngest child in the family has blue eyes, what is the probability that at least three of the children have blue eyes?
b. Explain why the answer in part (a) is different from the answer in Exercise 10.

### 6: 2.3.1

Suppose that $k$ events, $B_1, \dots, B_k$ form a partition of the sample space $S$. For $i = 1, \dots k$, let $P(B)$ denote the prior probability of $B_1$. Also, for each event $A$ such that $P(A) > 0$, let $P(B_i|A)$ denote the posterior probability of $B_i$ given that the event $A$ has occurred. Prove that if $P(B_1|A) < P(B_1)$, then $P(B_i|A) > P(B_i)$ for at least one value of $i$, $(i = 2, \dots k)$.

### 7: 2.3.4

A new test has been devised for detecting a particular type of cancer. If the test is applied to a person who has this type of cancer, the probability that the person will have a positive reaction is 0.95 and the probability that the person will have a negative reaction is 0.05. If the test is applied to a person who does not have this type of cancer, the probability that the person will have a positive reaction is 0.05 and the probability that the person will have a negative reaction is 0.95. Suppose that in the general population, one person out of every 100,000 people has this type of cancer. If a person selected at random has a positive reaction to the test, what is the probability that he has this type of cancer?

### 8: R - The Monte Hall Problem

is a famous problem in probability, and indeed, often very smart people are stumped.(See: Marilyn vos Savant, Ask Marilyn, Parade Magazine, 9 September; 2 December; 17 February 1990, reprinted in Marilyn vos Savant, Ask Marilyn, St. Martins, New York, 1992)

Suppose you're on Monty Hall's *Let's Make a Deal!* You are given the choice of three doors, behind one door is a car, the others, goats. You pick a door, say 1, Monty opens another door, say 3, which has a goat. Monty says to you "Do you want to pick door 2?" Is it to your advantage to switch your choice of doors?

you can play the game here: https://math.ucsd.edu/~crypto/Monty/monty.html

a. Write out the relevant events that need to be considered.

b. Provide the probabilities given in the problem (that is, report any marginal, conditional, and joint probabilities that are known/assumed).

c. What assumptions are needed to report the probabilities in b.?

d. Provide the probability statement relevant to the question at hand. Recall that the probability of interest has to do with switching doors. **Assume that the contestant chooses door 1 and Monty Hall opens door 3 without loss of generality.**

e. Using Bayes' Rule, find the probability of winning if the contestant switches doors (given they start with door 1 and switch to door 3).

f. Suppose that everything is the same except that Monty forgot to find out in advance which door has the car behind it. In the spirit of "the show must go on," he makes a guess at which of the two doors to open and gets lucky, opening a door behind which stands a goat. Now should the contestant switch? Again, **assume that the contestant chooses door 1 originally and is shown the goat behind door 3.**

g. For part f., what changed (as compared to the original game)?

h. Run an R simulation to estimate the probabilities computed in parts e. and f.

```
contestant = 1
car = sample(1:3,1)
goat = setdiff(1:3,car)
choicesE = setdiff(setdiff(1:3, car),contestant)
```

```r
choicesF = c(2,3)
contestant  # remove this line after you know what it does
```

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

```r
car         # remove this line after you know what it does
```

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

```r
goat        # remove this line after you know what it does
```

```
## [1] 2 3
```

```r
choicesE    # remove this line after you know what it does
```

```
## [1] 2 3
```

```r
choicesF    # remove this line after you know what it does
```

```
## [1] 2 3
```

```r
# setdiff does exactly what you think it'll do, a set difference!



if(length(choicesE) ==1){
  MH.choice = choicesE}
if(length(choicesE) > 1){
  MH.choice = sample(choicesE, 1)
}
MH.choice   # remove this line after you know what it does
```

```
## [1] 3
```

```r
# For part e.
n.reps=10
counter = 0
switch = c()
car.truth = c()
for (i in 1:n.reps){
  car = sample(1:3,1)
  goat = setdiff(1:3,car)
  choices = setdiff(setdiff(1:3, car),contestant)

  if(length(choices) ==1){
    MH.choice = choices}
  if(length(choices) > 1){
    MH.choice = sample(choices, 1)
  }

  car.truth = c(car.truth, car)
  switch = c(switch, setdiff(2:3, MH.choice))
}

table(switch, car.truth)
```

```
##       car.truth
## switch 1 2 3
##      2 2 3 0
```

```
##       3 1 0 4
```

```r
sum(switch==car.truth)/n.reps
```

```
## [1] 0.7
```

Hint for part f.: you need to change two things. 1. the choices from which Monty Hall picks the doors. 2. The way you tabulate the results. One easy way to tabulate only the simulated results of interest is to use an if statement around the part of the R code that tracks which door has the car and which door you'd switch to. Below is an example which is unrelated to the Monty Hall problem.

```r
# unrelated code which subsets the numbers less than or equal
# to 50 that are divisible by 5.

mynumbers = c()
for(i in 1:100){
  if(i <= 50 & (floor(i/5) == i/5)){
    mynumbers = c(mynumbers, i)
  }
}
mynumbers
```

```
##  [1]  5 10 15 20 25 30 35 40 45 50
```