

Note: “OIS” refers to OpenIntro Statistics, the textbook for this class.

Problem 1: OIS Exercise 3.8, page 93 (5 pts)

The American Community Survey is an ongoing survey that provides data every year to give communities the current information they need to plan investments and services. The 2010 American Community Survey estimates that 14.6% of Americans live below the poverty line, 20.7% speak a language other than English (foreign language) at home, and 4.2% fall into both categories.²²

- Are living below the poverty line and speaking a foreign language disjoint?
- (0 pts) Draw a Venn Diagram summarizing the variables and their associated probabilities. **This is not mandatory for the homework assignment! Do not include it in R Markdown.**
- What percent of Americans live below the poverty line and only speak English at home?
- What percent of Americans live below the poverty line or speak a foreign language at home?
- What percent of Americans live above the poverty line and only speak English at home?
- Is the event that someone lives below the poverty line independent of the event that the person speaks a foreign language at home?

Problem 2: OIS Exercise 3.11, page 94 (4 points)

The table below shows the distribution of education level attained by US residents by gender based on data collected in the 2010 American Community Survey (note that there are only two genders recorded).²³

	<i>Gender</i>	
	Male	Female
Less than 9th grade	0.07	0.13
9th to 12th grade, no diploma	0.10	0.09
HS graduate (or equivalent)	0.30	0.20
Some college, no degree	0.22	0.24
Associate's degree	0.06	0.08
Bachelor's degree	0.16	0.17
Graduate or professional degree	0.09	0.09
Total	1.00	1.00

- What is the probability that a randomly chosen man has at least a Bachelor's degree?
- What is the probability that a randomly chosen woman has at least a Bachelor's degree?
- What is the probability that a man and woman getting married both have at least a Bachelor's degree? Note any assumptions you must make to answer this question.
- If you made an assumption in part (c), do you think it was reasonable? If you didn't make an assumption, double check your earlier answer and then return to this part.

Problem 3: Conditional Probability in R (7 points)

Refer to the Titanic dataset we used in class.

```
titanic <- read.csv("http://vincentarelbundock.github.io/Rdatasets/csv/carData/TitanicSurvival.csv")
```

- Using `titanic`, create a table comparing the `survived` and `sex` variables (again, there are only two sexes recorded). Print the table. (1 pt)
- According to the table, what is the probability that a passenger survived? (1 pt)
- According to the table, what is the probability that a passenger was male? (1 pt)
- According to the table, what is the probability a passenger was male and survived? (1 pt)
- According to the table, what is the probability of a passenger surviving given that the passenger was male? (1 pt)
- Are sex and survival independent? Give a mathematical justification! (2 pts)

Problem 4 (4 pts)

Of volunteers coming into a blood center, 1 in 3 have type O^+ blood, 1 in 15 have O^- , 1 in 3 have A^+ , and 1 in 16 have A^- (*Note: These blood types are mutually exclusive— you cannot have both type O^- and A^+ at the same time*). This distribution can be summarized by the following table:

Blood Type	O^+	O^-	A^+	A^-
$p(\text{Blood Type})$	$\frac{1}{3}$	$\frac{1}{15}$	$\frac{1}{3}$	$\frac{1}{16}$

The name of one person who previously has donated blood is selected from the records of the center.

- Is this distribution described in the table a valid probability distribution? Why or why not?
- What is the probability that the randomly selected person has type O (either O^+ or O^-) blood?
- What is the probability that the randomly selected person has type A blood?
- What is the probability that the randomly selected person has neither type A or type O blood?

Problem 5 (5 pts)

A local college has three types of rooms in their dormitories: singles, doubles, and triples. Thus, the distribution of the number of roommates a randomly chosen student has, M , is as follows:

M	0	1	2
$p(M)$	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{3}{10}$

- In R, calculate the expected value of the number of roommates a randomly chosen student has. Save it as `exp_roommates`. (2)
- In R, calculate the variance of the number of roommates a randomly chosen student has. Save it as `var_roommates`. (3 pts)