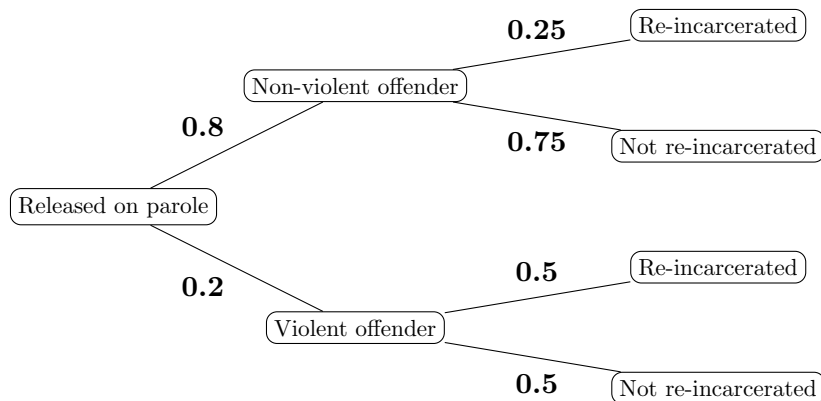


The following problems are similar to ones you might see on the midterm exam.

1. At one prison, 80% of prisoners released on parole had been convicted of non-violent offenses while the other 20% had been convicted of violent offenses. Suppose that 25% of non-violent paroles are re-incarcerated within five years while 50% of violent paroles are re-incarcerated within five years.

(a) Complete the tree diagram below by adding the correct weights to the branches.



(b) What percent of all paroles are re-incarcerated after five years?

**Solution:** There are two paths in the tree diagram that lead to a parole being re-incarcerated. A total of  $0.8 \cdot 0.25 = 0.2$  paroles are non-violent offenders and are re-incarcerated. Also,  $0.2 \cdot 0.5 = 0.1$  paroles are violent offenders and are re-incarcerated. So a total of  $0.3 = 30\%$  of paroles are re-incarcerated after 5 years.

(c) Given that a parole is not re-incarcerated after five years, what is the probability that they were originally convicted of a violent offense?

**Solution:** This is a conditional probability.

$$P(A | B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{0.2 \cdot 0.5}{0.7} = 14.3\%,$$

where  $A$  is the event that someone was a violent offender, and  $B$  is the event that they were not re-incarcerated after five years. To get  $P(A \text{ and } B) = 0.2 \cdot 0.5 = 0.1$ , just multiply along that path in the tree diagram.

2. A group of students are asked how many siblings they each have. The results are shown below.

# of Siblings	0	1	2	3	4	5
% of Students	20%	35%	25%	5%	10%	5%

What is the average number of siblings for these students?

**Solution:** This is a weighted average:

$$(0.20)(0) + (0.35)(1) + (0.25)(2) + (0.05)(2) + (0.10)(4) + (0.05)(5) = 1.6 \text{ siblings.}$$

3. The *Student Monitor* surveys 1200 undergraduates from 100 colleges semiannually to understand trends among college students. Recently, the *Student Monitor* reported that the average amount of time spent on the internet was 15.1 hours. Assume that the standard deviation is 5 hours.

- (a) Use the formula  $\bar{x} \pm 2\frac{s}{\sqrt{n}}$  to find a 95% confidence interval for the mean time spent per week on the internet.

**Solution:** The confidence interval is  $15.1 \pm 2\frac{5}{\sqrt{1200}} = 15.1 \pm 0.29$  which goes from 14.81 to 15.39 hours per week, on average.

- (b) What was the population in this study?
- A. The 1200 undergraduates.
  - B. Everyone in the United States.
  - C. All college students in the United States.**
  - D. The amount of time spent on the internet.
- (c) What was the variable in this study?
- A. The 1200 undergraduates.
  - B. Everyone in the United States.
  - C. All college students in the United States.
  - D. The amount of time spent on the internet.**
- (d) Which of the following best describes the confidence interval from part (a)?
- A. There is a 95% chance that the population mean is in the interval.**
  - B. There is a 95% chance that any college student is in the interval.
  - C. 95% of college students are in the interval.
  - D. There is a 95% chance that the average of a future survey will be in the interval.
4. A survey of 1124 likely voters in the Democratic primary in California before the 2016 election found that 584 planned to vote for Hillary Clinton.
- (a) Use the formula

$$\hat{p} \pm 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

to find a 95% confidence interval for the proportion of all Democratic primary voters who are planning to vote for Clinton.

**Solution:** The sample proportion is  $\hat{p} = \frac{584}{1124} = 0.52$  and the sample size is  $n = 1124$ , so the confidence interval is  $0.52 \pm 2\sqrt{\frac{0.52(0.48)}{1124}} = 0.52 \pm 0.03$ , which runs from 49% to 55%.

(b) What is the margin of error for this poll?

**Solution:** We already calculated the margin of error (it's the number after the  $\pm$  sign). It is 3%.

(c) What would decrease the margin of error for the poll?

- A. If we took a larger sample.
- B. If we took a smaller sample.
- C. If we took a random sample.
- D. If we carefully selected the sample to avoid bias.

5. The game of Texas hold'em starts with each player receiving two cards. Here is the probability model for the number of aces in two-card hands:

Number of aces	0	1	2
Probability	0.559	0.382	?

(a) What is the missing number in the table above?

**Solution:** The missing number is 0.059 (so that all three probabilities add up to exactly 1).

(b) What is the probability of getting at least one ace in a two-card hand?

**Solution:** Either one *or* two aces would count, so add the probabilities for those two outcomes: 44.1%.

6. The expected number of heads if you flip a fair coin once is 0.5, but of course you can't really get 0.5 heads. Give a one or two sentence explanation of what we mean by "expected value".

**Solution:** The expected value is the average outcome that you would expect if you repeated a random trial many many times. The law of large numbers says that the actual sample average tends to approach the expected value as the number of trials increases.

7. A 2010 Pew Research poll asked 1,306 Americans "From what you've read and heard, is there solid evidence that the average temperature on earth has been getting warmer over the past few decades, or not?". The two-table below shows the responses by party and ideology.

	Earth is warming	Not warming	Don't know	Total
Conservative Republican	144	261	26	431
Mod/Lib Republican	78	79	13	170
Mod/Cons Democrat	326	91	26	444
Liberal Democrat	235	13	13	261
Total	784	444	78	1306

- (a) What percent of all of the people in the sample were conservative Republicans?

**Solution:** This is a column proportion in the totals column. So it is  $\frac{431}{1306} = 33\%$ .

- (b) What percent of people who believe that the Earth is warming are conservative Republicans?

**Solution:** This is a row proportion focused on just conservative Republicans. It is  $\frac{144}{431} = 33.4\%$ .

- (c) How many times more likely are liberal Democrats to believe in global warming than conservative Republicans (i.e., what is the relative risk)?

**Solution:**  $\frac{235}{261} = 90.0\%$  of liberal Democrats believe in global warming, so they are  $\frac{90.0}{33.4} = 2.7$  times more likely to believe in global warming than conservative Republicans.

- (d) The row proportion  $326/444$  is a conditional probability, so it can be expressed using the notation  $P(A | B)$ . What is  $A$  and what is  $B$  for this particular conditional probability?

**Solution:** The row with 326 and 444 in it is the row for moderate/conservative Democrats. So the given is  $B$  = “moderate/conservative Democrat”. Within that row, 326 believe in global warming, so  $A$  = “believes in global warming”.

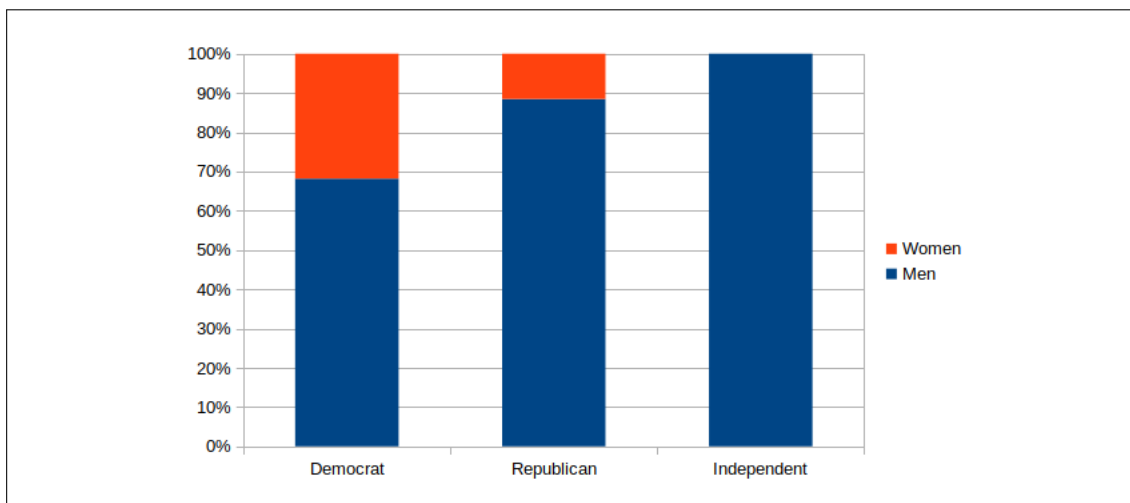
8. In the United States Senate in 2016, there were 44 Democrats, 54 Republicans, and 2 Independents. There were 20 women in the Senate (14 Democrats and 6 Republicans).

- (a) Make a **two-way table** and a **segmented bar graph** (with one bar for each party) to display the relationship between gender and political party in the Senate in 2016.

**Solution:** The two-way table is

	Democrat	Republican	Independent
Men	30	48	2
Women	14	6	0

A segmented bar graph showing the percents is



- (b) If you randomly picked a senator, what is the probability that you would pick a woman or a Democrat?

**Solution:** The probability of picking a woman is 20% and the probability of picking a Democrat is 44%. Using the addition rule, the probability of picking a woman or a Democrat is  $20\% + 44\% - 14\% = 50\%$ .

- (c) Were gender and political party associated or independent in the Senate?

**Solution:** Associated. Women were 31.8% of Democratic senators, but only 11.1% of Republican senators in 2016.