3.2: Marginal, Joint, and Conditional Probability

We have learned how probabilities can be estimated using relative frequencies from a table that separates data into groups according to the values of some variable. Similar methods can be used to investigate relationships between variables in a two-way table. We can compute various types of probabilities from a two-way table. In this section, we will explore **marginal, joint, and conditional probabilities**. We will do this using the data presented in the following two-way table.

The United States Census Bureau collects large amounts of data about the population of the United States.[3] The two-way table below presents marital status and whether an individual is male or female for the population of the United States. Numbers in this table are in millions of people. For example, 32.8 represents 32,800,000. Each value has been rounded to the nearest 0.1 million which is equivalent to 100,000.

Marital Status	Married	Widowed	Divorced	Separated	Never Married	Total
Male	64.5	3.4	12.2	2.1	47.5	129.7
Female	63.4	11.8	16.5	2.9	41.5	136.1
Total	127.9	15.2	28.7	5	89	265.8

Marginal Probability

If a probability is computed using only totals in the margins from the table (the far right column, or the bottom row in the above table), it is called a **marginal probability**. Let A represent a row or column heading in the table above. Then the marginal probability, $P(A)$, is given by

$$P(A) = \frac{\text{number of ways } A \text{ can happen}}{\text{total number of outcomes}}$$

In this fraction, the numerator (top of the fraction) is a row or column total from the margins of the table, and the denominator (bottom of the fraction) is the grand total (found in the lower right cell in the table above).

Use the two-way table above to find some probabilities:

1. What is the probability that a randomly selected person from the adult population in the US is male? In probability notation, you are asked to find $P(\text{male}) = \dfrac{\text{number of males}}{\text{grand total}}$. Round your answer to three decimal places.

2. What is the probability that a randomly selected person from the adult population in the US is divorced? Include probability notation in your answer. Round your answer to three decimal places.

   $P(\_\_\_\_\_\_\_) = \_\_\_\_\_\_\_$

Joint Probability

Frequently, we need to find a probability that involves two outcomes, both A and B. In a two-way table, probability of both events occurring, $P(AandB) = P(A \cap B)$, is found by taking the number of ways A and B can happen, and dividing by the grand total.

$$P(A \text{ and } B) = P(A \cap B) = \frac{\text{number of ways } A \text{ and } B \text{ can happen}}{\text{total number of outcomes}}$$

3. What is the probability that a randomly chosen adult is male and divorced? Round your answer to three decimal places.

   $P(\text{male} \cap \text{divorced}) = \_\_\_\_\_\_\_ =$

4. What is the proportion of adults that are female and never married? Include probability notation in your answer. Round your answer to three decimal places.

Multiplication Rule for Independent Events

When frequencies are given for values of two categorical variables in a two-way table, the joint probability can be computed without using any special rules. Without a two-way table, joint probability can sometimes be computed using the multiplication rule. If the events are **independent**, where the probability of one event does not depend on the occurrence of the other, then the probability can be computed using multiplication. If A and B are independent, formulaically, we can say

$$P(A \cap B) = P(A) \cdot P(B)$$

This is called the **multiplication rule for independent events.**

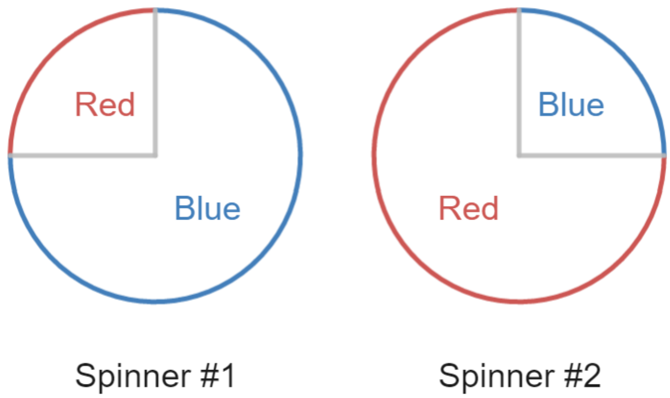
Consider a standard 52 deck of cards. What is the probability of selecting a queen and then an ace? The probability of selecting a queen is $\frac{4}{54}$ since there are 4 queens and 52 cards. This event changes the number of cards remaining in the deck. The probability of selecting an ace depends on these new frequencies. The probability of selecting an ace is now $\frac{4}{51}$ since there are 4 aces and 51 cards remaining in the deck. In this scenario, the events are dependent because **we do not return the chosen card(s) to the deck**. The events occurred **without replacement**.

$$P(\text{ queen } \cap \text{ ace }) = \frac{4}{52} \cdot \frac{4}{51} = \frac{4}{663} \approx 0.0060$$

Alternatively, if we return the chosen card(s) to the deck, the events are independent. In this scenario, the events occurred **with replacement**.

$$P(\text{ queen } \cap \text{ ace }) = \frac{4}{52} \cdot \frac{4}{52} \approx 0.0059$$

5. Consider a new spinner game:



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

a. You win if both spinners land on red. What is the probability that you win?

b. You win if both spinners land on blue. What is the probability that you win?

c. You win if the first spinner lands on red and the second spinner lands on blue. What is the probability that you win?

Conditional Probability

When two variables are **dependent**, the probability of one variable's outcome is influenced by the outcome of the other variable. If A and B are dependent, then the probability of B depends on a condition: whether A happens or not. We will calculate the probability of B given that A will occur. This is called **conditional probability**. The conditional probability of B, given that A has occurred is denoted $P(B \mid A)$ (read as "the probability of B given A").

6. Let's return to our census data from 2020. Let A represent widowed adults in the US, let B represent male adults in the US, and let C represent female adults in the US.

Marital Status	Married	$A =$ Widowed	Divorced	Separated	Never Married	Total
$B =$ Male	64.5	3.4	12.2	2.1	47.5	129.7
$C =$ Female	63.4	11.8	16.5	2.9	41.5	136.1
Total	127.9	15.2	28.7	5	89	265.8

a. Find the proportion of widowed individuals who are male. In other words, what is the probability that a randomly selected *widowed individual* is male? Now, we are not selecting from the entire population, but rather, from the subset that is defined by a condition. Therefore, the denominator of the fraction will be the total *widowed individuals* and the numerator will be the number of people who are *male and widowed*. Round your answer to three decimal places.

$$P(\text{male} \mid \text{widowed}) = P(B \mid A) = \frac{\text{number of male and widowed}}{\text{total number of widowed}} = \frac{P(A \cap B)}{P(A)} = \qquad \approx$$

b. Find the proportion of widowed individuals who are female. Include probability notation in your answer and round to three decimal places.

c. The number of widowed individuals who are male is close to the number of separated individuals who are male. Does this mean that the probabilities are the same? Explain.

The multiplication rule is altered for dependent events. If A and B are dependent (where the probability of one depends on the occurrence of the other), then $P(A \cap B) = P(A) \cdot P(B \mid A)$. Applying this rule, we get

$$P(A \cap B) = P(A) \cdot P(B \mid A) = P(\text{ widowed }) \cdot P(\text{ male } \mid \text{ widowed }) = \frac{15.2}{265.8} \cdot \frac{3.4}{15.2} = \frac{3.4}{265.8}$$
$$= \frac{\text{ways } A \text{ and } B \text{ can both happen}}{\text{total number of outcomes}}$$

This is the **multiplication rule for dependent events**. Remember that A and B is a single event in a two-way table, and when this is the case, the multiplication rule is unnecessary.

Reference

[3] https://data.census.gov/cedsci/table?q=Marital%20Status%20and%20Marital%20History&tid=ACSST5Y2020.S1201&moe=false