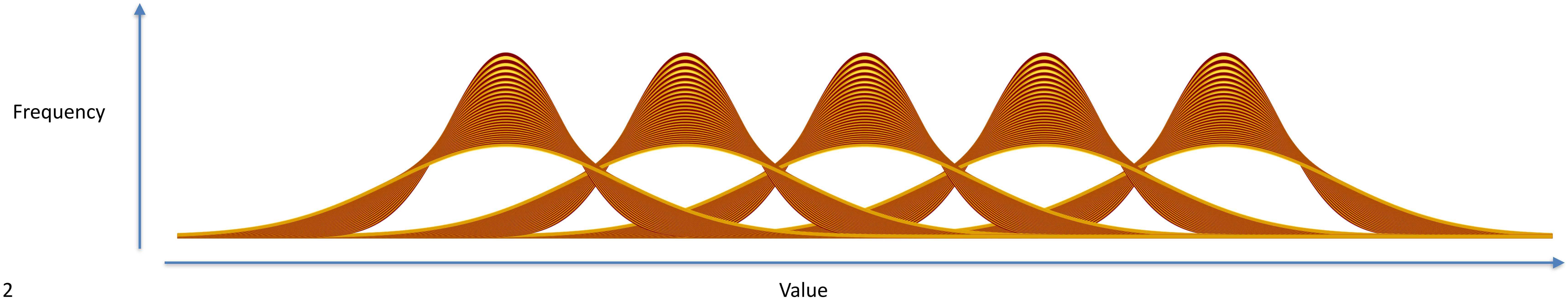
COLLEGE OF CHARLESTON

Week 3

Math 104: Elementary Statistics

Mathematics of statistics

- **probability**, the mathematics of chance
- **distribution**,
  - the frequency the various value of a variable occurs in a population
  - the probability of various values occurring
- p-values are key in inferential statistics
  - They are probabilities



4.1 Introduction to Probability

Probability: Definitions

- A **trial** is any process or experiment where the result is determined by chance
- An **outcome** is an individual result that is possible in a trial
- A **sample space** is the set of all outcomes in a trial
- An **event** is a subset of outcomes from sample space
- **Example**: tossing a coin with two sides (heads or tails) over and over
  - *trial* = tossing the coin
  - *outcome* = heads or tails (H or T)
  - *sample space* = {H, T}
  - event = particular subset of {H, T} which is one of the following: {H}, {T}, {H,T}, $\varnothing$

Probability: Definitions

- A **trial** is any process or experiment where the result is determined by chance
- An **outcome** is an individual result that is possible in a trial
- A **sample space** is the set of all outcomes in a trial
- An **event** is a subset of outcomes from sample space
- **Example**: rolling two 6-sided dice (each labeled 1 through 6) once
  - *trial* = rolling both dice
  - *outcome* = value of each dice (e.g. Dice 1 = 4, Dice 2 = 6)
  - *sample space* = all possible outcomes represented as orders pairs

$$\{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,1), (6,2), \ldots, (6,6)\}$$

  - *event* = subset of ordered pairs (e.g pairs that sum to 4)

$$(1,3), (2,2), (3,1)$$

Probability: Definitions

- A **trial** is any process or experiment where the result is determined by chance
- An **outcome** is an individual result that is possible in a trial
- A **sample space** is the set of all outcomes in a trial
- An **event** is a subset of outcomes from sample space
- **Example**: tossing a coin with two sides (heads or tails) once
  - *trial* = tossing the coin
  - *outcome* = sequence of heads and tails such as HHTTHTHHTTH
  - *sample space* = set of sequences of heads and tail
  - event = sequences of 5 trials that have at least 4 heads

Probability: Definitions

- The empirical or **experimental probability** is $\mathbb{P}(E) = \dfrac{f}{n}$ where E is an event, f is the frequency the outcome lies in the event, and n is the number of trials
- If every outcome is *equally likely*, the probability or **classical probability** is $\mathbb{P}(E) = \dfrac{n(E)}{n(S)}$ where $n(E)$ is the number of outcomes in the event and $n(S)$ is the number of outcomes in the sample space

Law of Large Numbers

- **Law of large numbers**: the greater the number of trials, the closer the experimental probability is to the classical probability.
- **Question**: if you repeatedly toss a coin, what do you expect to happen as you toss the coin more and more times?
  Answer: the empirical probability is very close to 0.5

Probability: Definitions

- The empirical or **experimental probability** is $\mathbb{P}(E) = \dfrac{f}{n}$ where E is an event, f is the frequency the outcome lies in the event, and n is the number of trials
- The probability or **classical probability** is $\mathbb{P}(E) = \dfrac{n(E)}{n(S)}$ where $n(E)$ is the number of outcomes in the event and $n(S)$ is the number of outcomes in the sample space
- **Example:** Beck is allergic to peanuts. At a large dinner party one evening, he notices that the cheesecake options on the dessert table contain the following flavors: 10 slices of chocolate, 12 slices of caramel, 12 slices of chocolate peanut butter, and 8 slices of strawberry. Assume that the desserts are served to guests at random.
  - What's the probability Beck's cheesecake contains peanuts?

    Answer: $2/7 \approx 0.2857$
  - What's the probability Beck's cheesecake does NOT contain chocolate?

    Answer: $10/21 \approx 0.4762$

•

Probability: Definitions

- The empirical or **experimental probability** is $\mathbb{P}(E) = \dfrac{f}{n}$ where E is an event, f is the frequency the outcome lies in the event, and n is the number of trials
- The probability or **classical probability** is $\mathbb{P}(E) = \dfrac{n(E)}{n(S)}$ where $n(E)$ is the number of outcomes in the event and $n(S)$ is the number of outcomes in the sample space
- **Example:** Consider a beginning archer who only manages to hit the target 50% of the time. What is the probability that in three shots, the archer will hit the target all three times?

  Answer: $0.5 \cdot 0.5 \cdot 0.5 = 0.125$

Probability: Definitions

- The empirical or **experimental probability** is $\mathbb{P}(E) = \dfrac{f}{n}$ where E is an event, f is the frequency the outcome lies in the event, and n is the number of trials
- The probability or **classical probability** is $\mathbb{P}(E) = \dfrac{n(E)}{n(S)}$ where $n(E)$ is the number of outcomes in the event and $n(S)$ is the number of outcomes in the sample space
- **Example:** Consider the novice archer again who hits the target 50% of the time. Assume that six shots have been taken and each time the target was hit. What is the probability that the next shot taken will also hit the target?

  Answer: Sample space $\{HHHHHHH, HHHHHHM\}$! Event $\{HHHHHHH\}$. So, $\dfrac{1}{2}$.

4.2 Addition Rules for Probability

Properties of Probability

- The probability is a number between 0 and 1
  - Why? $P(E) = \dfrac{n(E)}{n(S)}$ where $n(E)$ is the number of outcomes in an event and $n(S)$ is the number of outcomes in the sample space
  - An event that is very likely to happen is close to 1
  - An event that is nearly impossible to happen is close to 0
- Suppose S is a sample space and E is an event. Then,
  - For any event E, $0 \leq \mathbb{P}(E) \leq 1$
  - $P(S) = 1$
  - $P(\varnothing) = 0$

Complement Rule

- The **complement** of an event E denoted E^c is the set of all outcomes in the sample space S that are NOT in E.
  - $E + E^c = S$
- Suppose E is an event. Then, $\mathbb{P}(S) + \mathbb{P}(S^c) = 1$. Put another way, $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
- Thinking about complements:
  - What is the complement that none of a bunch of outcomes will happen after a series of trials?
- **Example**:
  - If you are worried that there is a 35% chance that you will fail your upcoming test, what is the probability that you will pass the test? Answer: 65%
  - If there is a 5% chance that none of the items on a scratch-off lottery ticket will be a winner, what is the probability that at least one of the scratch-of items will? Answer: 95%

Complement Rule

- The **complement** of an event E denoted E^c is the set of all outcomes in the sample space S that are NOT in E.
  - $E + E^c = S$
- Suppose E is an event. Then, $\mathbb{P}(S) + \mathbb{P}(S^c) = 1$. Put another way, $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
- Thinking about complements:
  - What is the complement that none of a bunch of outcomes will happen after a series of trials?
- **Example**:
  - Roll a pair of standard six-sided dice. What is the probability that neither die is a three? Answer: $1 - 11/36 = 25/36 \approx 0.6944$

Addition Rule

- Suppose E and F are two events. Then, $\mathbb{P}(E \texttt{ or } F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \texttt{ and } F)$
- **Mutually exclusive** events have no outcomes in common.
- If E and F are mutually exclusive, the $\mathbb{P}(E \texttt{ or } F) = \mathbb{P}(E) + \mathbb{P}(F)$
- **Example**: Suppose that after a vote in the US Senate on a proposed health care bill, the following table shows the breakdown of the votes by party. If a lobbyist stops a random senator after the vote, what is the probability that this senator will either be a Republican or have voted against the bill?

Answer: $\mathbb{P}(R \texttt{ or } A) = \mathbb{P}(R) + \mathbb{P}(A) - \mathbb{P}(R \texttt{ and } A) = \frac{50}{100} + \frac{32}{100} - \frac{7}{100} = \frac{75}{100}$

Votes on Health Care Bill		
	Voted in Favor	Voted Against
Democrat	23	21
Republican	43	7
Independent	2	4

Addition Rule

- Suppose E and F are two events. Then, $\mathbb{P}(E \texttt{ or } F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \texttt{ and } F)$
- **Mutually exclusive** events have no outcomes in common.
- If E and F are mutually exclusive, the $\mathbb{P}(E \texttt{ or } F) = \mathbb{P}(E) + \mathbb{P}(F)$
- **Example**: Roll a pair of dice. What is the probability of rolling either a total less than four or a total equal to ten?

  Answer: $\mathbb{P}(T < 4 \texttt{ or } T = 10) = \mathbb{P}(T < 4) + \mathbb{P}(T = 10) = \dfrac{3}{36} + \dfrac{3}{36} = \dfrac{1}{6}$

4.3 Multiplication Rules for Probability

Independence

- Two (or more!) events are **independent** if one event happening does not influence the probably of another. Otherwise, the events are **dependent**.
  - Independence is especially relevant in *multi-stage experiments* (e.g. repeatedly drawing balls in a lottery)
  - When objects are placed back into consideration before performing the next stage of a multistage experiment, we say the experiment is performed *with replacement*.
  - Otherwise, the experiment is performed *without replacement*
- Multiplication rule: suppose events E and F are independent. Then, $\mathbb{P}(E \texttt{ and } F) = \mathbb{P}(E) \cdot \mathbb{P}(F)/$
- **Example**: A booth at the state fair has rubber ducks floating in water. For each ticket, you get two chances to pick a duck from those floating by. In order to win a small stuffed animal, you must pick a duck that has a star drawn on the bottom. If you manage to choose two ducks that both have stars drawn on them, you win the extra-large stuffed animal. After your first pick, the rubber duck is placed back in the water for your second pick. If there are 50 floating ducks and 3 that have stars drawn on them, what is the probability of choosing 2 ducks that both have stars on them? Answer: $3/50 \cdot 3/50 = 9/2500$

Conditional Probability

- When calculating probability of dependent events, cannot apply previous rule!
- Conditional probability: $\mathbb{P}(F \,|\, E)$ = probability that F occurs given E
- Multiplication rule: $\mathbb{P}(E \texttt{ and } F) = \mathbb{P}(E) \cdot \mathbb{P}(F \,|\, E)$
- Conditional Formula: $\mathbb{P}(F \,|\, E) = \dfrac{\mathbb{P}(E \texttt{ and } F}{\mathbb{P}(E)}$
- **Example**: what is the probability of choosing two aces in a row from a standard deck of 52 cards? Assume there are 4 aces in a standard deck and that the cards are chosen without replacement.

  Answer: $\mathbb{P}(A1 \texttt{ and } A2) = \mathbb{P}(A1) \cdot \mathbb{P}(A2 \,|\, A1) = \dfrac{4}{52} \cdot \dfrac{3}{51} = \dfrac{12}{2652}$

Fundamental Counting Principle

- For a multistage experiment with n stages where ether first stage has k_1 outcomes, the second has k_2, …, the last stage has k_n outcomes, the total number of possible outcomes for a sequence of stages that make up the experiment are $k_1 \cdot k_2 \cdot \cdots \cdot k_n$.
- **Example**: a restaurant has a choice of 2 soups, 3 salads, 4 appetizers, 5 entrees, and 2 desserts. How many possible meals can you order at the restaurant (assuming you have all 5 courses)?

  Answer: $2 \cdot 3 \cdot 4 \cdot 5 \cdot 2 = 360$