Chapters 5-8

Stat 128: Elementary Statistics

A Document By:
Nathan Warner

July 03, 2023
Computer Science
Joliet Junior College
United States

Contents

Chapter 5

0.1 5.1: Probability Rules

Learning Objectives For This Section:

1. **Understand Random Processes and the Law of Large Numbers**

2. **Apply the Rules of Probabilities**

3. **Compute and Interpret Probabilities Using the Empirical Method**

4. **Compute and Interpret Probabilities Using the Classical Method**

5. **Recognize and Interpret Subjective Probabilities**

Vocab:

- **Simulation:** is a technique used to recreate a random event.

- **Random Process:** represents scenarios where the outcome of any particular trial of an experiment is unknown, but the proportion (or relative frequency) of a particular outcome is observed and approaches a specific value

- **Probability:** Is the measure of the likelihood of a random phenomenon or chance behavior occurring. It deals with experiments that yield random short-term results or outcomes yet reveal long-term predictability. The long-term proportion in which a certain outcome is observed is the probability of that outcome.

- **Outcomes:** Short term results

- **The Law of Large Numbers:** As the number of repetitions of a probability experiment increases, the proportion with which a certain outcome is observed gets closer to the probability of the outcome.

- **an experiment** is any process with uncertain results that can be repeated.

- **The sample space, S,** of a probability experiment is the collection of all possible outcomes for that experiment.

- **An event** is any collection of outcomes from a probability experiment. An event consists of one or more outcomes. We denote events with one outcome, sometimes called simple events, as e_i. In general, events are denoted using capital letters such as E.

- **A probability model** lists the possible outcomes of a probability experiment and each outcome's probability. A probability model must satisfy Rules 1 and 2 of the rules of probabilities.

- **An unusual event** is an event that has a low probability of occurring.

- An experiment has **equally likely outcomes** when each outcome has the same probability of occurring.

- A **subjective probability** is a probability that is determined based on personal judgment.

Formulas:

- Computing probability with the empirical method

$$P(E) = \text{Relative frequency of E} = \frac{Frequency\ of\ E}{number\ of\ trials\ of\ experiment}.$$

- Computing Probability With The Classical Method

  - If an experiment has n equally likely outcomes and if the number of ways that an event E can occur is m, then the probability of E,P(E), is

  $$P(E) = \frac{\text{number of ways that } E \text{ can occur}}{\text{number of possible outcomes}} = \frac{m}{n}.$$

  - So, if S is the sample space of this experiment, then

  $$P(E) = \frac{N(E)}{N(E)}.$$

  where N(E) is the number of outcomes in E, and N(S) is the number of outcomes in the sample space.

Objective 1: Understand Random Processes and the Law of Large Numbers:

Refer to vocab above for definition on the law of large numbers.

Question 1

Brad and Allison have three girls. Brad tells Allison that he would like one more child because they are due to have a boy. What do you think of Brad's logic?

Brad is incorrect due to the nonexistent Law of Averages. The fact that Brad and Allison had three girls in a row does not matter. The likelihood the next child will be a boy is about 0.5.

Example: A probability experiment consists of rolling a single six-sided fair die. A fair die is one in which each possible outcome is equally likely. For example, rolling a two is just as likely as rolling a five.

a.) Identify the outcomes of the probability experiment.

b.) Identify the sample space

c.) Define the event $E=$"roll an even number".

A.) The outcomes from rolling a single fair die are $e_1 =$ "rolling a one" $= \{1\}$, $e_2 =$ "rolling a two" $= \{2\}$, $e_3 =$ "rolling a three" $= \{3\}$, $e_4 =$ "rolling a four" $= \{4\}$, $e_5 =$ "rolling a five" $= \{5\}$, and $e_6 =$ "rolling a six" $= \{6\}$.

B.) The set of all possible outcomes forms the sample space, $S = \{1, 2, 3, 4, 5, 6\}$. There are six outcomes in the sample space.

C.) The event E="roll an even number"= $\{2, 4, 6\}$.

Question 2

If a person rolls a 6 sided dice and then flips a coin, describe the sample space of possible outcomes.

Solution:

The sample space is $\{1H, 1T, 2H, 2T, 3H, 3T, 4H, 4T, 5H, 5T, 6H, 6T\}$

Apply the Rules of Probabilities:

In the following probability rules, the notation $P(E)$ means "the probability that event E occurs."

Rules of Probability:

1. The probability of any event E, $P(E)$, must be greater than or equal to 0 and less than or equal to 1. That is, $0 \leqslant P(E) \leqslant 1$.

2. The sum of the probabilities of all outcomes must equal 1. That is, if the sample space $S = \{e_1, e_2, \ldots, e_n\}$, then

$$P(e_1) + P(e_2) + \cdots + P(e_n) = 1..$$

Example: Probability Model:

In a bag of peanut candies, the colors of the candies can be brown, yellow, red, blue, orange, or green. Suppose that a candy is randomly selected from a bag. The table shows each color and the probability of drawing that color. Verify that this is a probability model.

Color	Probability
Brown	0.12
Yellow	0.15
Red	0.12
Blue	0.23
Orange	0.23
Green	0.15

Solution:

Rule 1 is satisfied because all probabilities are between 0 and 1, inclusive.

Rule 2 is satisfied because $\sum P(E) = 1$

Key Concepts Regarding Probabilities

- If an event is impossible, the probability of the event is 0.

- If an event is a certainty, the probability of the event is 1.

- The closer a probability is to 1, the more likely the event will occur.

- The closer a probability is to 0, the less likely the event will occur.

- For example, an event with probability 0.8 is more likely to occur than an event with probability 0.75.

- An event with probability 0.8 will occur about 80 times out of 100 repetitions of the experiment, whereas an event with probability 0.75 will occur about 75 times out of 100.

Question 3

In a certain card game, the probability that a player is dealt a particular hand is 0.46 . Explain what this probability means. If you play this card game 100 times, will you be dealt this hand exactly 46 times? Why or why not?

Solution:

The probability 0.46 means that approximately 46 out of every 100 dealt hands will be that particular hand. No, you will not be dealt this hand exactly 46 times since the probability refers to what is expected in the long-term, not short-term.

Unusual Event

Typically, an event with a probability less than 0.05 (or 5%) is considered unusual, but this cutoff point is not set in stone. The researcher and the context of the problem determine the probability that separates unusual events from not so unusual events.

The three methods for determining the probability of an event:

- the Empirical Method
- the Classical Method
- the Subjective Method

Approximating Probabilities Using the Empirical Approach:

The probability of an event E occurring is approximately the number of times event E is observed divided by the number of repetitions (or trials) of the experiment.

$$P(E) = \text{Relative frequency of E} = \frac{Frequency\ of\ E}{number\ of\ trials\ of\ experiment}.$$

Note:-

When we find probabilities using the empirical approach, the result is approximate because different trials of the experiment lead to different outcomes and, therefore, different estimates of $P(E)$

Surveys are probability experiments. Why? Each time a survey is conducted, a different random sample of individuals is selected. Therefore, the results of a survey are likely to be different each time the survey is conducted because different people are included.

Compute and Interpret Probabilities Using the Classical Method

The empirical method gives an approximate probability of an event by conducting a probability experiment. The classical method of computing probabilities does not require that a probability experiment actually be performed. Rather, it relies on counting techniques.

The classical method of computing probabilities requires equally likely outcomes. An experiment has equally likely outcomes when each outcome has the same probability of occurring. For example, when a fair die is thrown once, each of the six outcomes in the sample space, 1,2,3,4,5,6, has an equal chance of occurring Computing Probability With The Classical Method

If an experiment has n equally likely outcomes and if the number of ways that an event E can occur is m, then the probability of E,P(E), is

$$P(E) = \frac{\text{number of ways that } E \text{ can occur}}{\text{number of possible outcomes}} = \frac{m}{n}.$$

So, if S is the sample space of this experiment, then

$$P(E) = \frac{N(E)}{N(E)}.$$

where N(E) is the number of outcomes in E, and N(S) is the number of outcomes in the sample space.

Comparing Empirical Probabilities and Classical Probabilities

We just saw that the classical probability of rolling a seven is 16 0.167. Suppose a pit boss at a casino rolls a pair of dice 100 times and obtains 15 sevens. From this empirical evidence, we would assign the probability of rolling a seven as 15100=0.15. If the dice are fair, we would expect the relative frequency of sevens to get closer to 0.167 as the number of rolls of the dice increases. In other words, the empirical probability will get closer to the classical probability as the number of trials of the experiment increases due to the Law of Large Numbers. If the two probabilities do not get closer, we may suspect that the dice are not fair.

> **Note:-**
>
> In simple random sampling, each individual has the same chance of being selected. Therefore, we can use the classical method to compute the probability of obtaining a specific sample.

Recognize and Interpret Subjective Probabilities

If a sports reporter is asked what he thinks the chances are for the Boston Red Sox to play in this season's World Series, the reporter would likely process information about the Red Sox (pitching staff, leadoff hitter, and so on) and then make an educated guess of the likelihood. The reporter may respond that there is a 20% chance the Red Sox will play in the World Series. This forecast is a probability, although it is not based on relative frequencies. We cannot, after all, repeat the experiment of playing a season under the same circumstances (same players, schedule, and so on) over and over. Nonetheless, the forecast of 20%=0.20 does satisfy the criterion that a probability be between 0 and 1, inclusive. This forecast is known as a subjective probability.

r Subjective probabilities are legitimate and are often the only method of assigning likelihood to an outcome. For instance, a financial reporter may ask an economist about the likelihood of the economy falling into recession next year. Again, we cannot conduct an experiment n times to find a relative frequency. The economist must use knowledge of the current conditions of the economy and make an educated guess about the likelihood of recession.

0.2 5.2: The Addition Rule and Complements

Learning Objectives For This Section:

1. **Use the Addition Rule for Disjoint Events**

2. **Use the General Addition Rule**

3. **Compute the Probability of an Event Using the Complement Rule**

Vocab:

- **Disjoint:** Two events are disjoint if they have no outcomes in common.

- **Mutually Exclusive:** Another name for disjoint events.

- **Complement of E:** Let S denote the sample space of a probability experiment and let E denote an event. The complement of E, denoted E^C, is all outcomes in the sample space S that are not outcomes in the event E.

Formulas:

- Addition Rule for Disjoint Events:

    - if E and F are disjoint (or mutually exclusive) events, then:

$$P(E \ or \ F) = P(E) + P(F).$$

- Complement Rule:

    - If E represents any event and E^C represents the complement of E, then

$$P(E^C) = 1 - P(E).$$

Use the Addition Rule for Disjoint Events

> #### _Definition: Addition Rule for Disjoint Events_
>
> If E and F are disjoint (or mutually exclusive) events, then:
>
> $$P(E \text{ or } F) = P(E) + P(F).$$

We often draw pictures of events using **Venn Diagrams**. These pictures represent events as circles enclosed in a rectangle. The rectangle represents the sample space, and each circle represents an event. For example, suppose we randomly select a chip from a bag where each chip in the bag is labeled 0,1,2,3,4,5,6,7,8,9. Let E represent the event "Choose a number less than or equal to 2", and let F represent the event "Choose a number greater than or equal to 8." These events are disjoint as shown in the figure.

Figure:

So you can see that these events are **Disjoint** because they **do not** share any outcomes.

We can then compute the probability of event E using the classical approach

$$P(E) = \frac{N(E)}{N(S)} = \frac{3}{10} = 0.3.$$

Similarly, if we wanted to know the probability of event F:

$$P(F) = \frac{N(F)}{N(S)} = \frac{2}{10} = 0.2.$$

What if we wanted to know the probability of E **or** F? ($P(E \text{ or } F)$):

$$P(E \text{ or } F) = \frac{N(E \text{ or } F)}{N(S)} = \frac{5}{10} = 0.5.$$

However, we can use the **Addition Rule**, which goes like:

$$P(E \text{ or } F) = P(E) + P(F) = 0.3 + 0.2$$
$$= 0.5.$$

Compute the Probability of an Event Using the Complement Rule

Suppose that the probability of an event E is known and we would like to determine the probability that E does not occur. This can be accomplished using the idea of complements.

Because E and E^C are mutually exclusive,

> ### *Definition:*
>
> **Complement Rule:**
>
> - If E represents any event and E^C represents the complement of E, then
>
> $$P(E^C) = 1 - P(E).$$

Example: Compliment Rule

Suppose 31.6% of American households own a dog. What is the probability that a randomly selected household **does** not own a dog?

$$\begin{aligned} P(E^C) &= 1 - P(E) \\ &= 1 - 0.36 \\ &= 0.684. \end{aligned}$$

Example: Compliment Rule:

Let:

$$\begin{aligned} S &= \{9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20\} \\ E &= \{11, 12, 13, 14, 15, 16, 17, 18, 19\}. \end{aligned}$$

Assume each outcome is equally likely. List the outcomes in E^C Find $P(E^C)$

$$\begin{aligned} P(E) &= \frac{9}{12} = \frac{3}{4} = 0.75 \\ P(E^C) &= 1 - 0.75 \\ &= 0.25. \end{aligned}$$

0.3   5.3: Independence and the Multiplication Rule

Learning Objectives For This Section:

1. **Identify Independent Events**

2. **Compute At-least Probabilities**

3. **Use the Multiplication Rule for Independent Events**

Vocab:

- **Two events being Independent:** Two events E and F are independent if the occurrence of event E in a probability experiment does not affect the probability of event F.

- **Two events being Dependent:** Two events are dependent if the occurrence of event E in a probability experiment affects the probability of event F.

Formulas:

- If E and F are independent events, then

$$P(E \text{ and } F) = P(E) \cdot P(F).$$

- If $E_1, E_2, E_3, \ldots, E_n$ are independent events, then

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } \ldots \text{ and } E_n) = P(E_1) \cdot P(E_2) \cdot \ldots \cdot P(E_n).$$

Identify Independent Events

Example A: Suppose you draw a card from a standard 52-card deck of cards and then roll a die. The events "draw a heart" and "roll a number" are **independent** because the results of choosing a card do not impact the results of the die toss.

Example B: Suppose two 40-year old women who live in the United States are randomly selected. The events "woman 1 survived the year" and "women 2 survives the year" are **independent**

Example C: Suppose two 40-year old women live in the same apartment complex. The events "woman 1 survived the year" and "women 2 survives the year" are **dependent**

Note:-

When we take small samples from large finite populations, we make the assumption of independence even though the events are technically dependent.

As a general rule of thumb, if the sample size n is no more than 5% of the population N ($n \leqslant 0.05N$), we assume independence.

Question 4

Determine whether the events E and F are independent or dependent. Justify your answer.

E: A person having an at-fault accident
F: The same person being prone to road rage

Solution: E and F are dependent because being prone to road rage can affect the probability of a person having an at-fault accident

E: A randomly selected person coloring her hair black.
F: A different randomly selected person coloring her hair blond.

Solution: E cannot affect F and vice versa because the people were randomly selected, so the events are independent.

E: The rapid spread of a cocoa plant disease
F: The price of chocolate

Solution: The rapid spread of a cocoa plant disease could affect the price of chocolate, so E and F are dependent.

Disjoint (or Mutually Exclusive) Events versus Independent Events

Disjoint events and independent events are different concepts.

- Recall that two events are disjoint if they have no outcomes in common, that is, if knowing that one of the events occurs, we know that the other event did not occur.

- Independence means that one event occurring does not affect the probability of the other event occurring.

Therefore, knowing that two events are disjoint means that the events are not independent.

Consider the experiment of rolling a single die. Let E represent the event "roll an even number" and let F represent the event "roll an odd number." We can see that E and F are disjoint (mutually exclusive) because they have no outcomes in common. In addition, $P(E) = \frac{1}{2}$ and $P(F) = \frac{1}{2}$. However, if we are told that the roll of the die is going to be an even number, then what is the probability of event F? Because the outcome will be even, the probability of event F is now 0 (and the probability of event E is now 1). So knowledge of event E changes the likelihood of observing event F.

Use the Multiplication Rule for Independent Events

Suppose that you flip a fair coin twice. What is the probability that you obtain a head on both flips, that is, a head on the first flip and a head on the second flip? If H represents the outcome "heads" and T represents the outcome "tails," then the sample space of this experiment is

$$S = HH, HT, TH, TT.$$

There is one outcome with both heads. Because each outcome is equally likely, we have

$$P(\text{heads on Flip 1 and heads on Flip 2}) = \frac{N(\text{heads on Flip 1 and heads on Flip 2})}{N(S)}$$
$$= \frac{1}{4}.$$

We may have intuitively figured this out by recognizing $P(\text{head}) = \frac{1}{2}$ for each flip. So it seems reasonable that

$$P(\text{heads on Flip 1 and heads on Flip 2}) = P(\text{heads on Flip 1}) \cdot P(\text{heads on Flip 2})$$
$$= \frac{1}{2} \cdot \frac{1}{2}$$
$$= \frac{1}{4}.$$

Because both approaches result in the same answer, we conjecture that $P(E \text{ and } F) = P(E) \cdot P(F)$, which is true.

Definition: Multiplication Rule for Independent Events

If E and F are independent events, then

$$P(E \text{ and } F) = P(E) \cdot P(F).$$

If $E_1, E_2, E_3, \ldots, E_n$ are independent events, then

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } \ldots \text{ and } E_n) = P(E_1) \cdot P(E_2) \cdot \ldots \cdot P(E_n).$$

Question 5

Suppose that events E F are independent, $P(E) = 0.6$ and $P(F) = 0.7$ What is the $P(E \text{ and } F)$?

$$P(E \text{ and } F) = P(E) \cdot P(F) = 0.6 \cdot 0.7 = 0.42.$$

Compute At-least Probabilities

Usually, when computing probabilities involving the phrase at least, use the Complement Rule.

The phrase at least means "greater than or equal to." For example, a person must be at least 17 years old to see an R-rated movie. This means that the person's age must be greater than or equal to 17 to watch the movie.

Question 6

The probability that a randomly selected female aged 60 years will survive the year is 0.99186 according to the National Vital Statistics Report. What is the probability that at least one of 500 randomly selected 60-year-old females will die during the course of the year?

Solution:

The phrase at least means "greater than or equal to," so we want to know the probability that 1 or 2 or 3 or or 500 60-year-old females will die during the year. These events are mutually exclusive, so

$$P(1 \text{ or } 2 \text{ or } 3 \text{ or } \ldots \text{ or } 500 \text{ die}) = P(1 \text{ dies}) + P(2 \text{ dies}) + P(3 \text{ dies}) + \cdots + P(500 \text{ dies}).$$

Computing these probabilities would be very time-consuming. However, notice that the complement of "at least one dying" is "none die", or all 500 survive. Use the Complement Rule to compute the probability.

So:

$$\begin{aligned} P(\text{at least one dies}) &= 1 - P(\text{None dies}) \\ &= 1 - (0.99186)^{500} \\ &= 0.9832. \end{aligned}$$

Chapter 6

6.1: Discrete Random Variables

Learning Objectives For This Section:

1. **Distinguish between Discrete and Continuous Random Variables**

2. **Identify Discrete Probability Distributions**

3. **Graph Discrete Probability Distributions**

4. **Compute and Interpret the Mean of a Discrete Random Variable**

5. **Interpret the Mean of a Discrete Random Variable as an Expected Value**

6. **Compute the Standard Deviation of a Discrete Random Variable**

Vocab:

- A **random variable** is a numerical measure of the outcome of a probability experiment; so its value is determined by chance.

    - Random variables are typically denoted using capital letters such as X

- A **discrete random variable** has either a finite or countable number of values. The values of a discrete random variable can be plotted on a number line with space between each point. See Figure 1(a).

- A **continuous random variable** has infinitely many values. The values of a continuous random variable can be plotted on a line in an uninterrupted fashion. See Figure 1(b).

- The **probability distribution** of a discrete random variable X provides the possible values of the random variable and their corresponding probabilities. A probability distribution can be in the form of a table, graph, or mathematical formula.

- Because the mean of a random variable represents what we would expect to happen in the long run, it is also called the **expected value**, $E(X)$. The interpretation of the expected value is the same as the interpretation of the mean of a discrete random variable.

Formulas:

- **Notation for probability of discrete random variables**

$$P(X).$$

  - Read as:

    "The probability that the random variable X equals x"

- **Mean of a Discrete Random Variable**

$$\mu_X = \sum_{i=1}^{n} x \cdot P(x).$$

  Where x is the value of the random variable and $P(x)$ is the probability of observing the value x.

- **Expected Value:**

$$E(X) = \mu_X.$$

- **Standard Deviation of a Discrete Random Variable**

$$\sigma_X = \sqrt{\sum [(x - \mu_X)^2 \cdot P(x)]}.$$

  Where x is the value of the random variable, μ_X is the mean of the random variable, and $P(x)$ is the probability of observing x.

- **Variance of discrete random variable:**

$$\mu^2.$$

Distinguish between Discrete and Continuous Random Variables

(a) The number of A's earned in a section of statistics with 15 students enrolled is a discrete random variable because its value results from counting. If the random variable X represents the number of A's, then the possible values of X are $x = 0, 1, 2, \ldots, 15$.

(b) The number of cars that travel through a McDonald's drive-through in the next hour is a discrete random variable because its value results from counting. If the random variable X represents the number of cars, the possible values of X are $x = 0, 1, 2, \ldots$. That is, the number of cars can be any whole number, and we do not impose an upper limit on the number of cars.

(c) The speed of the next car that passes a state trooper is a continuous random variable because speed is measured. If the random variable S represents the speed, the possible values of S are all positive real numbers; that is, $s > 0$. Even though a radar gun may report the speed of a car as 37 miles per hour, it is actually any number greater than or equal to 36.5 mph and less than 37.5 mph. That is, $36.5 \leqslant s < 37.5$.

Identify Discrete Probability Distributions

Rules for a Discrete Probability Distribution:

Let $P(x)$ denote the probability that the random variable X equals x; then

- $\sum P(x) = 1$
- $0 \leqslant P(x) \leqslant 1$

Graph Discrete Probability Distributions

In the graph of a discrete probability distribution, the horizontal axis is the value of the discrete random variable and the vertical axis is the corresponding probability of the discrete random variable. When graphing a discrete probability distribution, we want to emphasize that the data are discrete. Therefore, draw the graph of discrete probability distributions using vertical lines above each value of the random variable to a height that is the probability of the random variable.

Discrete Probability Distribution Figure:

Distribution Shape of Discrete Probability Distributions

Graphs of discrete probability distributions help determine the shape of the distribution.

Recall that we describe distributions as skewed left, skewed right, or symmetric. The graph in Figure 2 is skewed left.

Compute and Interpret the Mean of a Discrete Random Variable

Finding Mean of a Discrete Random Variable In Statcrunch:

1. Stat > Calculators > Custom

2. Select Values in:

3. Select Weights in (P(X)):

4. Compute

How to Interpret the Mean of a Discrete Random Variable

The mean of a discrete random variable can be thought of as the mean outcome of the probability experiment if we repeated the experiment many times.

As the number of repetitions of the experiment increases... The mean value of the n trials will approach μ_X (The difference between $\bar{x}$ and μ_X gets closer to 0 as n increases)

Interpret the Mean of a Discrete Random Variable as an Expected Value

Example: A term life insurance policy will pay a beneficiary a certain sum of money upon the death of the policy holder. These policies have premiums that must be paid annually. Suppose a life insurance company sells a \$250,000 one-year term life insurance policy to a 49-year-old female for \$530. According to the National Vital Statistics Report, Vol. 47, No. 28, the probability that the female will survive the year is 0.99791. Compute the expected value of this policy to the insurance company.

Approach:

The experiment has two possible outcomes: survival or death. Let the random variable X represent the payout (money lost or gained), depending on survival or death of the insured. Assign probabilities to each payout and substitute these values into $\mu_X = \sum (x \cdot P(X))$

Step 1. Because $P(\text{survives}) = 0.99791$, $P(\text{dies}) = 0.00209$. If the client survives the year, the insurance company makes \$530, or $x = +530$. If the client dies during the year, the insurance company must pay \$250,000 to the client's beneficiary, but still keeps the \$530 premium; so $x = \$530 - \$250,000 = -\$249,470$. The value is negative because it is money paid by the insurance company. The probability distribution is listed in Table 3.

x	$P(X)$
\$530 (survives)	0.99791
-\$249,470 (dies)	0.00209

Step 2. The expected value (from the point of view of the insurance company) of the policy is

$$E(X) = \mu_X$$
$$= \sum (x \cdot P(X))$$
$$= (\$530 \cdot 0.99791) + (-\$249470 \cdot 0.00209)$$
$$= \$7.50.$$

Therefore:

The company expects to make \$7.50 for each 49-year-old female client it insures. The \$7.50 profit of the insurance company is a long-term result. It does not make \$7.50 on each 49-year-old female it insures; rather, the average profit per 49-year-old female insured is \$7.50. Because this is a long-term result, the insurance "idea" will not work with only a few insured.

Compute the Standard Deviation of a Discrete Random Variable

Statcrunch Steps:

1. Stat > Calculators > Custom

2. Select Values in:

3. Select Weights in (P(X)):

4. Compute

6.2: The Binomial Probability Distribution

Learning Objectives For This Section:

1. **Determine Whether a Probability Experiment is a Binomial Experiment**

2. **Compute Probabilities of Binomial Experiments**

3. **Compute the Mean and Standard Deviation of a Binomial Random Variable**

4. **Graph a Binomial Probability Distribution**

Vocab:

- The **binomial probability distribution** is a discrete probability distribution that describes probabilities for experiments in which there are two mutually exclusive (disjoint) outcomes. These two outcomes are generally referred to as success (such as making a free throw) and failure (such as missing a free throw). Experiments in which only two outcomes are possible are referred to as binomial experiments, provided that certain criteria are met.

- **A combination** is a collection, without regard to order and without repetition, in which r objects are chosen from n distinct objects with $r \leqslant n$.

- **Trial:** Each repetition of the experiment.

Formulas:

- **nCk (Combinations):**

    - The n objects are distinct
    - Repetition of objects is not allowed, and
    - Order is not important

$$nCk = \frac{n!}{(k!(n-k)!)}.$$

- **Binomial Probability Distribution Function**

$$P(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

    where p is the probability of success.

- **Mean (or Expected Value) and Standard Deviation of a Binomial Random Variable**

    - A binomial experiment with n independent trials and probability of success p has mean, μ_X, and standard deviation, σ_X, given by the formulas

$$\mu_X = np, \text{ and standard deviation, } \sigma_X = \sqrt{np(1-p)}.$$

Determine Whether a Probability Experiment Is a Binomial Experiment

Critera:

An experiment is said to be a **binomial experiment** if:

1. The experiment is preformed a fixed number of times. Each repetition of the experiment is called a **trial**

2. The trials are independent. This means that the outcome of one trial will not affect the outcome of the other trials.

3. for each trial, there are two mutually exclusive outcomes, success or failure.

4. The probability of success is fixed for each trial of the experiment.

Notation used in the binomial probability distribution:

- There are n independent trials of the experiment.

- let p denote the probability of success so that $1 - p$ is the probability of failure

- Let X be a binomial random variable that denotes the number of successes in n independent trials of the experiment. So $0 \leqslant X \leqslant n$

Compute Probabilities of Binomial Experiments

Now we will compute probabilities for a binomial random variable X. We present three methods for obtaining binomial probabilities:

1. The binomial probability distribution formula

2. A table of binomial probabilities

3. Technology

Using the Binomial Probability Distribution Function in Statcrunch

Steps:

1. Stat > Calculators > Binomial

2. Enter n

3. Enter p

4. Enter info for $P(X)$

5. Compute

Compute the Mean and Standard Deviation of a Binomial Random Variable

Example: According to CTIA, 55% of all U.S. households are wireless-only households. In a simple random sample of 500 households, determine the mean and standard deviation number of wireless-only households. Compute the mean and std. dev

Approach:

This is a binomial experiment with $n = 500$ and $p = 0.55$. Using $\mu_X = np$, we can find the mean as $\mu_x = \sqrt{np(1-p)}$

Solution:

$$\mu_X = np = 500(0.55) = 275$$
$$\text{and}$$
$$\sigma_X = \sqrt{np(1-p)} = \sqrt{500(0.55)(1 - 0.55)} = \sqrt{123.75}$$
$$\approx 11.1.$$

Graph a Binomial Probability Distribution

To graph a binomial probability distribution, first find the probabilities for each possible value of the random variable. Then follow the same approach as was used to graph discrete probability distributions.

Shape of Binomial Probability Distribution for Various Values of p

The binomial probability distribution is

- skewed right if $p < 0.5$

- symmetric and approximately bell-shaped if $p = 0.5$

- skewed left of $p > 0.5$

Shape of the Graph of a Binomial Probability Distribution for Various Values of n

- For a fixed p, as the number of trials n in a binomial experiment increases, the probability distribution of the random variable X becomes bell-shaped.

- As a rule of thumb, if $np(1p) \geqslant 10$, the probability distribution will be approximately bell-shaped.

> **Note:-**
>
> This result allows us to use the Empirical Rule to identify unusual observations in a binomial experiment. Recall that the Empirical Rule states that in a bell-shaped distribution, about 95% of all observations lie within 2 standard deviations of the mean. That is, about 95% of the observations lie between -2 and $+2$. Any observation that lies outside this interval may be considered unusual because the observation occurs less than 5% of the time.

Example: According to CTIA, 55% of all U.S. households are wireless-only. In a simple random sample of 500 households, 301 were wireless-only. Is this result unusual?

Approach:

Because $np(1-p) = 500(0.55)(1-0.55) = 123.75 \geqslant 10$, the binomial probability distribution is approximately bell-shaped. Therefore, we can use the Empirical Rule: If the observation is less than $\mu - 2\sigma$ or greater than $\mu + 2\sigma$, it is unusual.

Solution: We have $\mu_X = 500(0.55) = 275$ and $\sigma_X = \sqrt{np(1-p)} = \sqrt{500(0.55)(1-0.55)} = 11.1$. Now,

$$\mu_X - 2\sigma_X = 275 - 2(11.1) = 275 - 22.2 = 252.8$$
$$and$$
$$\mu_X + 2\sigma_X = 275 + 2(11.1) = 275 + 22.2 = 297.2..$$

Any value less than 252.8 or greater than 297.2 is unusual; therefore, 301 is an unusual result. We should try to identify the reason for its value. Perhaps the percentage of households that are wireless-only has increased.

Chapter 7

7.1: Properties of the Normal Distribution

Learning Objectives For This Section:

1. **Use the Uniform Probability Distribution**

2. **Graph a Normal Curve**

3. **State the Properties of the Normal Curve**

4. **Explain the Role of Area in the Normal Density Function**

Vocab:

- **Probability Density Function (pdf):** an equation used to compute probabilities of continuous random variables. It must satisfy the following two properties:

  1. The total area under the graph of the equation over all possible values of the random variable must equal 1

  2. The height of the graph of the equation must be greater than or equal to 0 for all possible values of the random variable. That is, the graph of the equation must lie on or above the horizontal axis for all possible values of the random variable.

- The **area under the graph of the density function** over an interval represents the probability of observing a value of the random variable in that interval.

- In mathematics, a **model** is an equation, a table, or a graph used to describe reality.

- The magenta curve in Figure 1 is a model called the **normal curve**, which is used to describe continuous random variables that are normally distributed.

Figure 1:
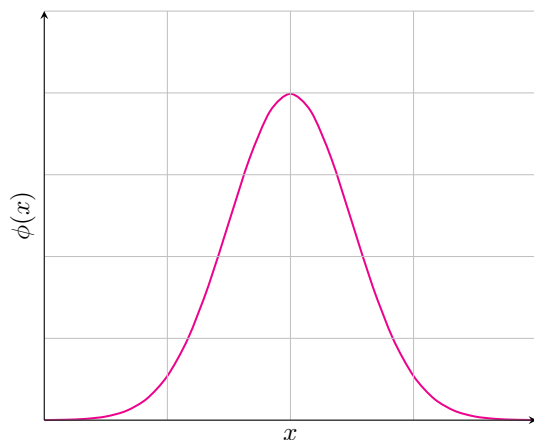
- A continuous random variable is **normally distributed**, or has a **normal probability distribution**, if its relative frequency histogram has the shape of a normal curve.

- **Inflection Points:** On a bell shaped curve (normal curve), the graph has two inflection points. At $\mu \pm \sigma$, these points are where the graph shifts from increasing at an increasing rate, to increasing at a decreasing rate

Use the Uniform Probability Distribution:

Assume that United Parcel Service is supposed to deliver a package to your front door and the arrival time is somewhere between 10 AM and 11 AM. Let the random variable X represent the time from 10 AM when the delivery is supposed to take place.

The delivery could be at 10 AM ($x = 0$) or at 11 AM ($x = 60$), with all one-minute intervals of time between $x = 0$ and $x = 60$ equally likely. That is to say, your package is just as likely to arrive between 10:15 and 10:16 as it is to arrive between 10:40 and 10:41.

The random variable X can be any value in the interval from 0 to 60, that is, $0 \leqslant X \leqslant 60$. Because any two intervals of equal length between 0 and 60, inclusive, are equally likely, the random variable X is said to follow a **uniform probability distribution**.

When computing probabilities for discrete random variables, we usually substitute the value of the random variable into a formula.

Things are not as easy for continuous random variables. Because an infinite number of outcomes are possible for continuous random variables, the probability of observing one particular value is zero. In the UPS example, the probability that the package arrives exactly 12.9438823 minutes after 10 AM is zero. This result is based on classical probability: there is one way to observe 12.9438823, and there are an infinite number of possible values between 0 and 60. To resolve this problem, we compute probabilities of continuous random variables over an interval of values. For example, we might compute the probability that your package arrives between $x = 10$ minutes and $x = 15$ minutes after 10 AM. To find probabilities for continuous random variables, we use probability density functions.

Uniform Density Function for UPS example:

The graph illustrates the properties for the "time" example. Notice the area of the rectangle is one and the graph is greater than or equal to zero for all x between 0 and 60, inclusive

Because the area of a rectangle is $l \cdot h$, and the width of the rectangle is 60, the height must be $\frac{1}{60}$

> **Note:-**
>
> Values of the random variable X less than 0 or greater than 60 are impossible, thus the equation must be zero for X less than 0 or greater than 60.

Area as a Probability:

The probability of choosing a time that is between 15 and 30 seconds after the minute is the area under the uniform density function.

Graph a Normal Curve

When describing a uniform random variable using a probability distribution, a rectangle is used to find probabilities of observing an interval of numbers (such as 10 to 20 minutes after 10 AM). However, not all continuous random variables follow a uniform distribution. For example, continuous random variables such as IQ scores and birth weights of babies have distributions that are symmetric and bell-shaped. Consider the histograms in Figure 1, which represent the IQ scores of 10,000 randomly selected adults. Notice that as the class width of the histogram decreases, the histogram becomes closely approximated by the smooth red curve. For this reason, we can use the curve to model the probability distribution of this continuous random variable.

Figure:

State the Properties of the Normal Curve:

The normal probability density function satisfies all the requirements of probability distributions. Now we list the properties of the normal density curve. The video explains the properties.

1. Because mean = median = mode, the normal curve has a single peak and the highest point occurs at $x = \mu$.

2. The normal curve has inflection points at $\mu - \sigma$ and $\mu + \sigma$.

3. The area under the normal curve is 1.

4. The area under the normal curve to the right of μ equals the area under the normal curve to the left of μ, which equals 0.5.

5. As x increases without bound (gets larger and larger), the graph approaches, but never reaches, the horizontal axis. As x decreases without bound (gets more and more negative), the graph approaches, but never reaches, the horizontal axis.

6. The Empirical Rule:

   - Approximately 68% of the area under the normal curve is between $x = \mu - \sigma$ and $x = \mu + \sigma$.
   - Approximately 95% of the area is between $x = \mu - 2\sigma$ and $x = \mu + 2\sigma$.
   - Approximately 99.7% of the area is between $x = \mu - 3\sigma$ and $x = \mu + 3\sigma$.

Explain the Role of Area in the Normal Density Function:

Area under a Normal Curve

Suppose that a random variable X is normally distributed with mean μ and standard deviation σ. The area under the normal curve for any interval of values of the random variable X represents either

- The proportion of the population with the characteristic described by the interval of values,

  or

- The probability that a randomly selected individual from the population will have the characteristic described by the interval of values.

7.2: Applications of the Normal Distribution

Learning Objectives For This Section:

1. **Find and Interpret the Area under a Normal Curve**

2. **Find the Value of a Normal Random Variable**

Vocab:

- If a normal random variable X has a mean different from 0 or a standard deviation different from 1, we can transform X into a **standard normal random variable Z** whose mean is 0 and standard deviation is 1.

Find and Interpret the Area under a Normal Curve

We use z-scores to find the area under a normal curve by hand. Recall that the z-score allows us to transform a random variable X with mean μ and standard deviation σ into a random variable Z with mean 0 and standard deviation 1.

Standardizing a Normal Random Variable:

Suppose that the random variable X is normally distributed with mean μ and standard deviation σ.

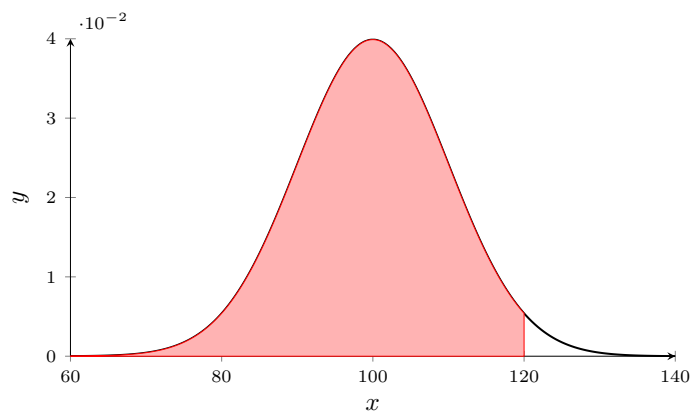Then the random variable

$$Z = \frac{X - \mu}{\sigma}.$$

is normally distributed with mean $\mu = 0$ and standard deviation $\sigma = 1$.

The random variable Z is said to have the **standard normal distribution.**

Consider an example:

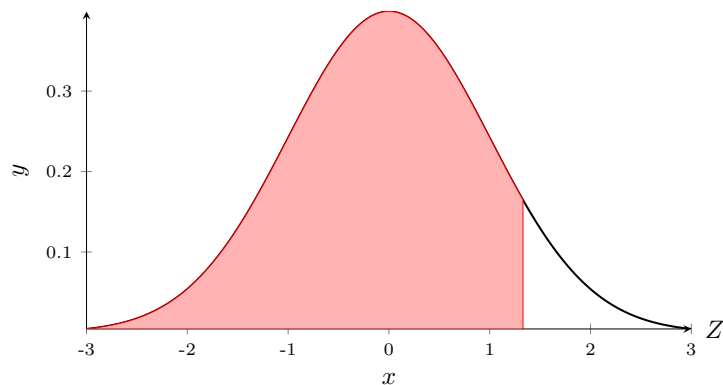IQ scores can be modeled by a normal distribution with $\mu = 100$ and $\sigma = 15$ Suppose a person has an IQ of 120

Figure:

Now we can transform this into a standard normal (Z distribution)

$$Z = \frac{120 - 100}{15}$$
$$= 1.33.$$

Figure:

Then by looking at the Table, we can determine the area in the red shaded region. Which is .9082. From this we can also determine that the area in the red region of our original distribution is also .9082

Finding an interpreting Area under a Normal Curve With Statcrunch:

1. Stat > Calculators > Normal

2. Input Mean

3. Input Standard Deviation

4. Input X

5. Compute

Find the Value of a Normal Random Variable:

Often, we do not want to find the proportion, probability, or percentile given a value of a normal random variable. Rather, we want to find the value of a normal random variable that corresponds to a certain proportion, probability, or percentile. For example, we might want to know the height of a 3-year-old girl who is at the 20th percentile. Or we might want to know the scores on a standardized exam that separate the middle 90% of scores from the bottom and top 5%.

Example: The heights of a pediatrician's 3-year-old females are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Find the height of a 3-year-old female at the 20th percentile.

Approach:

Use .2 in Statcrunch steps mentioned previously as what P(X) is equal to.

Example: The mean incubation time of fertilized eggs is days. Suppose the incubation times are approximately normally distributed with a standard deviation of day. Determine the incubation times that make up the middle 97%

Solution:

So in statcrunch, we will use values 0.05, and 0.97

Important Notation For the Future:

The notation z_α (pronounced "z sub alpha") is the z-score such that the area under the standard normal curve to the right of z_α is α.

Example: Find the value of $Z_{0.21}$

Solution: So we want to find the area to the right of 0.21, in statcrunch:

1. Stat > Calculator > Normal

2. Mean: 0

3. Std. Dev: 1

4. $P(x \geqslant 0) = 0.21$

5. Compute!

7.3 Assessing Normality

Learning Objectives For This Section:

1. **Use Normal Probability Plots to Assess Normality**

Vocab:

- A **normal probability plot** is a graph that plots observed data versus normal scores.

- A **normal score** is the expected z-score of the data value, assuming that the distribution of the random variable is normal. The expected z-score of an observed value depends on the number of observations in the data set.

Use Normal Probability Plots to Assess Normality:

Up to this point, we have said that a random variable X is normally distributed, or at least approximately normal, provided the histogram of the data is symmetric and bell-shaped. This works well for large data sets, but the shape of a histogram drawn from a small sample of observations does not always accurately represent the shape of the population. For this reason, we need additional methods for assessing the normality of a random variable X when we are looking at a small set of sample data.

Drawing a Normal Probability Plot by Hand:

1. Arrange the data in ascending order.

2. Compute:

$$f_i = \frac{i - 0.375}{n + 0.25}.$$

   where i is the index (the position of the data value in the ordered list) and n is the number of observations. The expected proportion of observations less than or equal to the ith data value is f_i.

3. Find the z-score corresponding to f_i from Table V.

4. Plot the observed values on the horizontal axis and the corresponding expected z-scores on the vertical axis.

The idea behind finding the expected z-score is, if the data come from a normally distributed population, we could predict the area to the left of each data value. The value of f_i represents the expected area to the left of the ith observation when the data come from a population that is normally distributed. For example, f_1 is the expected area to the left of the smallest data value, f_2 is the expected area to the left of the second-smallest data value, and so on.

Values of normal random variables and their z-scores are linearly related ($x = \mu + z\sigma$), so a plot of observations of normal variables against their expected z-scores will be linear. We conclude the following:

If sample data are taken from a population that is normally distributed, a normal probability plot of the observed values versus the expected z-scores will be approximately linear.

It is difficult to determine whether a normal probability plot is "linear enough." However, we can use a procedure based on the research of S. W. Looney and T. R. Gulledge in their paper "Use of the Correlation Coefficient with Normal Probability Plots," published in the American Statistician. Basically, if the linear correlation coefficient between the observed values and expected z-scores is greater than the critical value found in Table VI, then it is reasonable to conclude that the data could come from a population that is normally distributed.

Normal probability plots are typically drawn using technology (graphing calculators or statistical software). However, it is worthwhile to go through an example that demonstrates how to draw a normal probability plot by hand to better understand the results supplied by technology.

Drawing a Normal Probability Plot Using Statcrunch

1. Graph > QQ plot

2. Select Column

3. Select Correlation Statistic

4. Compute

Chapter 8

8.1 Distribution of the Sample Mean

Learning Objectives For This Section:

1. **Describe the Distribution of the Sample Mean: Normal Population**

2. **Describe the Distribution of the Sample Mean: Non-normal Population**

Vocab:

- The **sampling distribution** of a statistic is a probability distribution for all possible values of the statistic computed from a sample of size n

- The **sampling distribution of the sample mean** $\overline{x}$ is the probability distribution of all possible values of the random variable $\overline{x}$ computed from a sample of size n from a population with mean μ and standard deviation σ.

- The standard deviation of the sampling distribution of $\overline{x}$, $\sigma_{\overline{x}}$, is called the **standard error** of the mean. **Note:** As the sample size n increases, the standard error of the mean **decreases**

- **The Central Limit Theorem:** Regardless of the shape of the underlying population, the sampling distribution of $\overline{x}$ becomes approximately normal as the sample size, n, increases

Formulas:

- **Mean of sampling distribution**

$$\mu_{\overline{x}} = \mu.$$

- **Standard Deviation of sampling distribution**

$$\sigma_{\overline{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}}.$$

- **finite population correction factor**

$$\sqrt{\frac{N-n}{N-1}}.$$

Introduction: The idea that the sample mean is a random variable

Suppose the government wanted to determine the mean income of all U.S. households. One approach the government could take is literally to survey each U.S. household to determine the population mean. This would be a very expensive and time-consuming survey.

A second approach the government could (and does) take is to survey a random sample of U.S. households and use the results to estimate the mean income of all households. The Current Population Survey is administered to approximately 60,000 randomly selected households each month. Among the many questions on the survey, respondents are asked to report the income of each individual in the household. From this information, the federal government obtains a sample mean U.S. household income. For example, in 2015, the mean annual U.S. household income was approximately $\bar{x}=\$79{,}263$. The government might infer from this survey that the mean annual household income of all U.S. households in 2015 was $\$\mu=79{,}263$.

The households in the Current Population Survey were determined by chance (random sampling). A second random sample of households would likely lead to a different sample mean, such as $\bar{x}=\$78{,}518$, and a third random sample of households would likely lead to a third sample mean, such as $\bar{x}=\$79{,}081$. Because the households selected will vary from sample to sample, the sample mean of household income will also vary from sample to sample. For this reason, **the sample mean is a random variable**; so it has a probability distribution. Our goal in this section is to describe the distribution of the sample mean. Remember, when we describe a distribution, we do so in terms of its shape, center, and spread.

Steps for Determining the Sampling Distribution of the Sample Mean

1. Obtain a simple random sample of size n

2. Compute the sample mean.

3. Assuming that we are sampling from a finite population, repeat Steps 1 and 2 until all distinct simple random samples of size n have been obtained.

> **Note:-**
>
> Once a particular sample is obtained, it cannot be obtained a second time.

Describe the Distribution of the Sample Mean: Normal Population

Key points:

- **Shape:** the shape of the distribution of the sample mean becomes approximately normal as the sample size n increases, regardless of the shape of the underlying population

- **Center:** The mean of the distribution of the sample mean will equal the mean of the parent population, $\mu_{\bar{x}} = \mu$.

- **Spread:** The standard devitaino of the distribution of the sample mean is less than the standard deviation of the population. And, the larger the sample size, n, the smaller the standard deviation of the distribution of the sample mean.

we conclude that the mean of the distribution of the sample mean, $\bar{x}$ equals the mean of the parent population and as the sample size n increases, the standard deviation of the distribution of $\bar{x}$ decreases. Although the proof is beyond the scope of this course.

The Mean and Standard Deviation of the Sampling Distribution of $\bar{x}$:

Suppose that a simple random sample of size n is drawn from a population with mean μ and standard deviation σ. The sampling distribution of $\bar{x}$ has mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. The standard deviation of the sampling distribution of $\bar{x}$, $\sigma_{\bar{x}}$, is called the standard error of the mean.

Note: Technically, we assume that we are drawing a simple random sample from an infinite population. For populations of finite size N,

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}}.$$

However, if the sample size is less than 5% of the population size $n < 0.05N$, the effect of $\sqrt{\frac{N-n}{N-1}}$ (the finite population correction factor) can be ignored without significantly affecting the result.

In summary:

Say we have a set of data, Where each row represents a sample of the population, we can compute $\bar{x}$ for each row of data, and then with this new row of $\bar{x}$'s, we can compute $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$

Now that we can find the mean and standard deviation for any sampling distribution of x¯¯¯, we can concentrate on the shape of the distribution. In the simulations from the video, both histograms of the 2000 different sample means appear to be normal. Recall that the population from which the sample was drawn was also normal. This result leads us to believe that if the population is normal, then the distribution of the sample mean is also normal.

The Shape of the Sampling Distribution of $\bar{x}$ If X Is Normal

If a random variable X is normally distributed, the distribution of the sample mean, $\overline{x}$, is normally distributed.

Example: The IQ, X, of humans is approximately normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$. Compute the probability that a simple random sample of size $n = 10$ results in a sample mean greater than 110. That is, compute $P(\bar{x} > 110)$.

Approach

The random variable X is normally distributed, so the sampling distribution of $\bar{x}$ will also be normally distributed. Verify the sample size is less than 5% of the population size (the independence requirement). The mean of the sampling distribution is $\mu_{\bar{x}} = \mu$, and its standard deviation is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. To find the probability by hand, convert the sample mean $\bar{x} = 110$ to a z-score and then find the area under the standard normal curve to the right of this z-score, or use technology to find the area.

Solution:

So we find $\mu_{\overline{x}}$ and $\sigma_{\overline{x}}$:

$$\mu_{\overline{x}} = \mu = 100$$

$$\sigma_{\overline{x}} = \frac{s}{\sqrt{n}} = \frac{15}{\sqrt{10}} = 4.7434.$$

Now in stat crunch:

- Stat > Calculators > Normal

- Input $\mu_{\overline{x}}$ as the mean

- Input $\sigma_{\overline{x}}$ as the standard deviation

- Input $P(x \geqslant 110)$

- Compute

And we end up with

.0.175.

Describe the Distribution of the Sample Mean: Non-normal Population

1. **Shape:** The shape of the distribution of the sample mean becomes approximately normal as the sample size n increases, regardless of the shape of the underlying population.

2. **Center:** The mean of the distribution of the sample mean will equal the mean of the parent population, $\mu_{\overline{x}} = \mu$

3. **Spread:** The standard deviation of the distribution of the sample mean is less than the standard deviation of the population, And, the larger the sample size, n, the smaller the standard deviation of the distribution of the sample mean.

There are two key concepts to understand from the above points:

1. The mean of the sampling distribution of the sample mean is equal to the mean of the underlying population. That is, $\mu_{\overline{x}=\mu}$ regardless of the size of the sample. The standard deviation of the sampling distribution of the sample mean is $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$, regardless of the size of the sample.

2. The shape of the distribution of the sample mean becomes approximately normal as the sample size n increases, regardless of the shape of the underlying population. We formally state point 2 as the Central Limit Theorem.