Name____

Probability Distributions

A random variable is a variable whose value is determined by the outcomes of a random event. For example, when you roll a six-sided die, you can define a random variable X that represents the number showing on the die. So, the possible values of X are 1, 2, 3, 4, 5, and 6. For every random variable, a *probability distribution* can be defined.

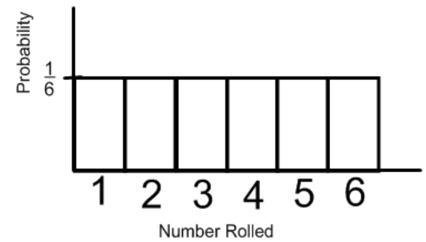
A histogram of this table would look like this:

Probability Distributions

A **probability distribution** is a function that gives the probability of each possible value of a random variable. The sum of all the probabilities in a probability distribution must equal 1.

Probability Distribution for Rolling a Die

X	1	2	3	4	5	6
***P*(*X*)**	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

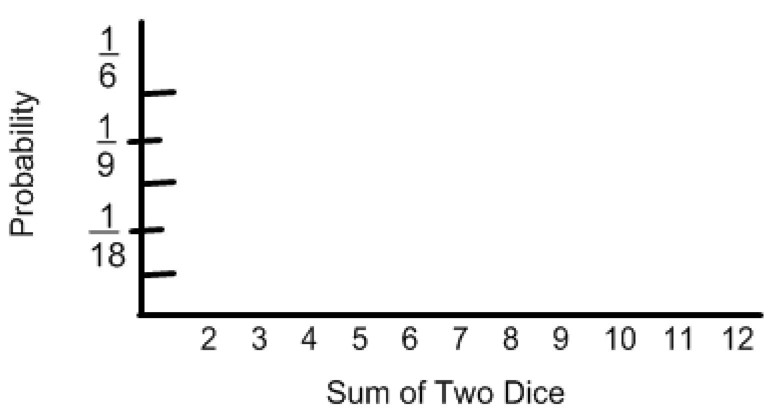


1. Two 6-sided dice are tossed. Let X be a random variable that represents the sum of the two dice. Complete the table showing the sum of the two dice.

Sum	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Make a table and a histogram showing the probability distribution for X.

X(sum)	2	3	4	5	6	7	8	9	10	11	12
# outcomes											
P(X)											



a. what is the most likely sum when rolling two six-sided dice?

b. What is the probability that the sum of the two dice is at least 10?

2. Four coins are flipped. Complete the tree/table showing the possible outcomes.

<table>
<tr><th>Coin 1</th><th>Coin 2</th><th>Coin 3</th><th>Coin 4</th><th>Result</th></tr>
<tr><td rowspan="8">H</td><td rowspan="4">H</td><td rowspan="2">H</td><td>H</td><td>HHHH</td></tr>
<tr><td>T</td><td>HHHT</td></tr>
<tr><td rowspan="2">T</td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="4">T</td><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="8">T</td><td rowspan="4"></td><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="4"></td><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
<tr><td rowspan="2"></td><td></td><td></td></tr>
<tr><td></td><td></td></tr>
</table>

Complete the probability distribution table and make a histogram.

# Heads	0	1	2	3	4
# outcomes					
P(X)					

Histogram: (Be sure to accurately number and LABEL.)

a. What is the most likely outcome of tossing 4 coins?

b. What is the probability that at least 2 heads are tossed?

3. Let X be a random variable that represents the sum when two four-sided dice are rolled. Complete the table showing the sums possible. Make a table and histogram showing the probability distribution for x.

Sum	1	2	3	4
1				
2				
3				
4				

X(sum)	2	3	4	5	6	7	8
# outcomes							
P(X)							

Histogram: (Be sure to accurately number and LABEL.)

a. What is the most likely outcome of rolling the two dice?

b. What is the probability that the sum of the two dice is <u>at most</u> 4? Explain/show work.

4. Let X be the letter on a block chosen randomly from a bag. The bag contains 7 blocks labeled "A.", 3 blocks labeled "B," 6 blocks labeled "C," and 5 blocks labeled "D."

X(letter)	**A**	**B**	**C**	**D**
# outcomes				
P(X)				

Histogram: (Be sure to accurately number and LABEL.)

BINOMIAL DISTRIBUTIONS One type of probability distribution is a **binomial distribution**. A binomial distribution shows the probabilities of the outcomes of a *binomial experiment.*

Binomial Experiments

A **binomial experiment** meets the following conditions:

- There are n independent trials.
- Each trial has only two possible outcomes: success and failure.
- The probability of success is the same for each trial. This probability is denoted by p. The probability of failure is given by $1 - p$.

For a binomial experiment, the probability of exactly k successes in n trials is:

$$P(k \text{ successes}) = {}_nC_k p^k (1 - p)^{n-k}$$

Construct a binomial distribution

SPORTS SURVEYS According to a survey, about 41% of U.S. households have a soccer ball. Suppose you ask 6 randomly chosen U.S. households whether they have a soccer ball. Draw a histogram of the binomial distribution for your survey.

Solution

The probability that a randomly selected household has a soccer ball is $p = 0.41$. Because you survey 6 households, $n = 6$.

AVOID ERRORS
You can check your calculations for a binomial distribution by adding all the probabilities. The sum should always be 1.

$P(k = 0) = {}_6C_0(0.41)^0(0.59)^6 \approx 0.042$

$P(k = 1) = {}_6C_1(0.41)^1(0.59)^5 \approx 0.176$

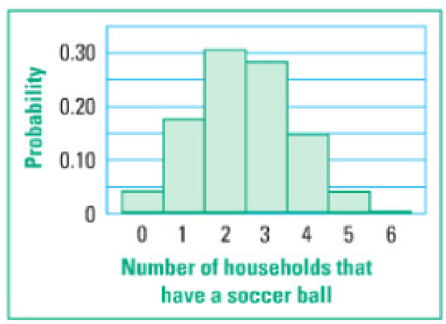
$P(k = 2) = {}_6C_2(0.41)^2(0.59)^4 \approx 0.306$

$P(k = 3) = {}_6C_3(0.41)^3(0.59)^3 \approx 0.283$

$P(k = 4) = {}_6C_4(0.41)^4(0.59)^2 \approx 0.148$

$P(k = 5) = {}_6C_5(0.41)^5(0.59)^1 \approx 0.041$

$P(k = 6) = {}_6C_6(0.41)^6(0.59)^0 \approx 0.005$



A histogram of the distribution is shown.

Interpret a binomial distribution

Use the binomial distribution in Example 3 to answer each question.

a. What is the most likely outcome of the survey?

b. What is the probability that at most 2 households have a soccer ball?

Solution

a. The most likely outcome of the survey is the value of k for which $P(k)$ is greatest. This probability is greatest for $k = 2$. So, the most likely outcome is that 2 of the 6 households have a soccer ball.

b. The probability that at most 2 households have a soccer ball is:

$$P(k \le 2) = P(k = 2) + P(k = 1) + P(k = 0)$$
$$\approx 0.306 + 0.176 + 0.042$$
$$\approx 0.524$$

▶ So, the probability is about 52%.

5. About 25% of adults in the U.S. are myopic (near-sighted.) Suppose you randomly survey 7 adults. Complete the probability calculations below and draw a histogram of the binomial distribution for your survey.

The probability that a randomly selected adult is myopic is p = _____. The probability that they are not myopic is 1 – p = _____

Because you survey 7 adults, n = _____.

P(k = 0) = ____________________________ ≈ _____

Draw histogram below:

P(k = 1) = ____________________________ ≈ _____

P(k = 2) = ____________________________ ≈ _____

P(k = 3) = ____________________________ ≈ _____

P(k = 4) = ____________________________ ≈ _____

P(k = 5) = ____________________________ ≈ _____

P(k = 6) = ____________________________ ≈ _____

P(k = 7) = ____________________________ ≈ _____

a. What is the most likely outcome of the survey?

b. What is the probability that at least 3 of the people are near-sighted?

6. A survey taken at your school found that 68% of the students are not afraid to fly. Suppose you randomly survey 5 students. Complete the probability calculations below and draw a histogram of the binomial distribution for your survey.

The probability that a randomly selected student is not afraid to fly is p = _____.

The probability that they <u>are</u> afraid to fly is 1 – p = _____

Because you survey 5 students, n = _____.

P(k = 0) = ____________________________ ≈ _____

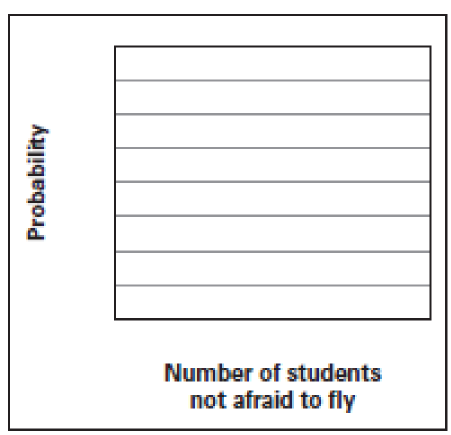
P(k = 1) = ____________________________ ≈ _____

P(k = 2) = ____________________________ ≈ _____

P(k = 3) = ____________________________ ≈ _____

P(k = 4) = ____________________________ ≈ _____

P(k = 5) = ____________________________ ≈ _____



a. What is the most likely outcome of the survey?

b. What is the probability that at least 3 students are not afraid to fly?

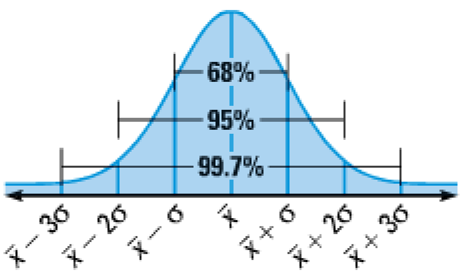
Part 2 - Normal Distributions

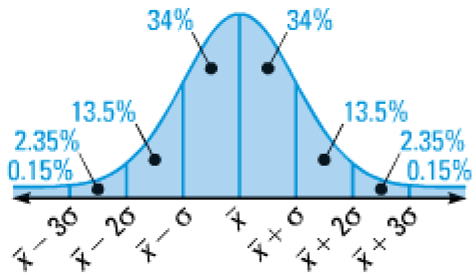
One type of probability distribution is a *normal distribution.* A **normal distribution** is modeled by a bell-shaped curve called a **normal curve** that is symmetric about the mean.

Areas Under a Normal Curve

A normal distribution with mean $\overline{x}$ and standard deviation σ has the following properties:

- The total area under the related normal curve is 1.
- About 68% of the area lies within 1 standard deviation of the mean.
- About 95% of the area lies within 2 standard deviations of the mean.
- About 99.7% of the area lies within 3 standard deviations of the mean.





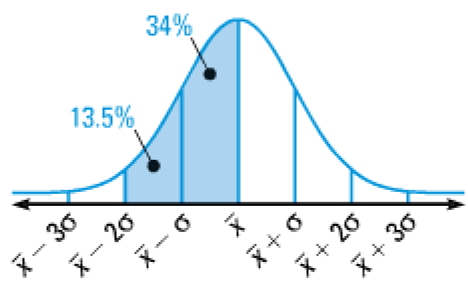
Example:

A normal distribution has mean $\overline{x}$ and standard deviation σ. For a randomly selected x-value from the distribution, find $P(\overline{x} - 2\sigma \le x \le \overline{x})$.

Solution

The probability that a randomly selected x-value lies between between $\overline{x} - 2\sigma$ and $\overline{x}$ is the shaded area under the normal curve shown.

$P(\overline{x} - 2\sigma \le x \le \overline{x}) = 0.135 + 0.34 = 0.475$



A normal distribution has mean $\overline{x}$ and standard deviation σ. Using the normal curves above, find the indicated probability for a randomly selected *X*-value from the distribution.

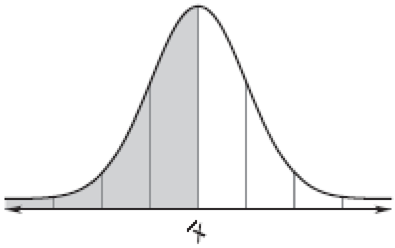
1. $P(x \ge \overline{x})$

2. $P(x \le \overline{x} - 2\sigma)$

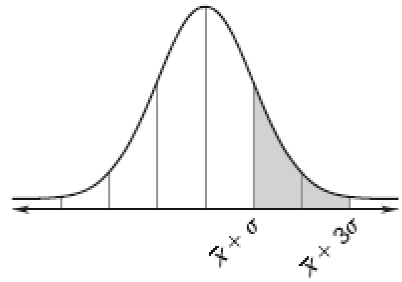
3. $P(x \le \overline{x} + 3\sigma)$

Give the percent of the area under the normal curve represented by the shaded region.

4.



5.



<u>Example :</u>

Interpret normally distributed data

READING
The abbreviation "mg/dl" stands for "milligrams per deciliter."

USE PERCENTILES
When n% of the data are less than or equal to a certain value, that value is called the *nth percentile*. Part (b) of Example 2 shows that 158 is the 16th percentile. Similarly, 172 is the 50th percentile.

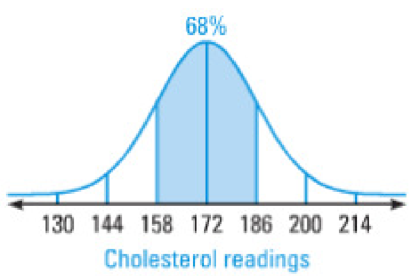
HEALTH The blood cholesterol readings for a group of women are normally distributed with a mean of 172 mg/dl and a standard deviation of 14 mg/dl.

a. About what percent of the women have readings between 158 and 186?

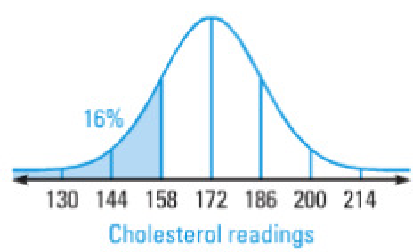
b. Readings less than 158 are considered desirable. About what percent of the readings are desirable?

Solution

a. The readings of 158 and 186 represent one standard deviation on either side of the mean, as shown below. So, 68% of the women have readings between 158 and 186.



b. A reading of 158 is one standard deviation to the left of the mean, as shown. So, the percent of readings that are desirable is 0.15% + 2.35% + 13.5%, or 16%.



A normal distribution has mean $\bar{x}$ and standard deviation σ. Find the indicated probability for a randomly selected x-value from the distribution.

1. $P(x \le \bar{x})$

2. $P(x \ge \bar{x})$

3. $P(\bar{x} \le x \le \bar{x} + 2\sigma)$

4. $P(\bar{x} - \sigma \le x \le \bar{x})$

5. $P(x \le \bar{x} - 3\sigma)$

6. $P(x \ge \bar{x} + \sigma)$

7. WHAT IF? In Example 2, what percent of the women have readings between 172 and 200?

A normal distribution has a mean of 18 and a standard deviation of 3. Find the probability that a randomly selected -*X*-value from the distribution is in the given interval.

8. Between 18 and 21

9. Between 12 and 18

10. Between 15 and 24

11. At least 21

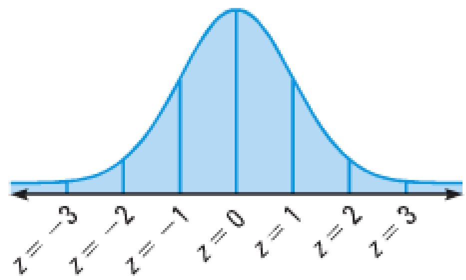
12. At least 27

13. At most 12

STANDARD NORMAL DISTRIBUTION The **standard normal distribution** is the normal distribution with mean 0 and standard deviation 1. The formula below can be used to transform *x*-values from a normal distribution with mean $\overline{x}$ and standard deviation σ into *z*-values having a standard normal distribution.

Formula: $z = \frac{x - \overline{x}}{\sigma}$

Subtract the mean from the given *x*-value, then divide by the standard deviation.



The *z*-value for a particular *x*-value is called the **z-score** for the *x*-value and is the number of standard deviations the *x*-value lies above or below the mean $\overline{x}$.

z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−3	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000+
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981
3	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.0000−

Example:

BIOLOGY Scientists conducted aerial surveys of a seal sanctuary and recorded the number *x* of seals they observed during each survey. The numbers of seals observed were normally distributed with a mean of 73 seals and a standard deviation of 14.1 seals. Find the probability that at most 50 seals were observed during a survey.

Solution

STEP 1 **Find** the *z*-score corresponding to an *x*-value of 50.

$$z = \frac{x - \overline{x}}{\sigma} = \frac{50 - 73}{14.1} \approx -1.6$$

STEP 2 **Use** the table to find $P(x \le 50) \approx P(z \le -1.6)$.

The table shows that $P(z \le -1.6) = 0.0548$. So, the probability that at most 50 seals were observed during a survey is about 0.0548.

z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−3	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000+
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159

A normal distribution has a mean of 50 and a standard deviation of 5. Use the standard normal table above to find the indicated probability for a randomly selected -*X*-value from the distribution.

14. $P(x \leq 50)$

15. $P(x \leq 55)$

16. $P(x \leq 40)$

17. $P(x \leq 62)$

18. $P(x \leq 47)$

19. $P(x \leq 34)$

20. $P(x \geq 48)$

21. $P(x \geq 54)$

22. $P(x \geq 39)$

22. $P(x \geq 63)$

22. $P(x \geq 45)$

23. $P(x \geq 36$

In Exercises 24 and 25, use the following information.

Restaurant Seating A restaurant is busiest Saturday from 5:00 P.M. to 8:00 P.M. During these hours, the waiting time for customers in groups of 4 or less to be seated is normally distributed with a mean of 15 minutes and a standard deviation of 2 minutes.

24. What is the probability that customers in groups of 4 or less will wait 17 minutes or less to be seated during the busy Saturday night hours?

25. What is the probability that customers in groups of 4 or less will wait 17 minutes or more to be seated during the busy Saturday night hours?

In Exercises 26 and 27, use the following information.

Light Bulbs A company produces light bulbs having a life expectancy that is normally distributed with a mean of 2000 hours and a standard deviation of 50 hours.

26. Find the *z*-score for a life expectancy of 2085 hours.

27. What is the probability that a randomly selected light bulb will last at most 2085 hours?

28. The time a fire department takes to arrive at the scene of an emergency is normally distributed with a mean of 6 minutes and a standard deviation of 1 minute.

a. What is the probability that the fire department takes at most 8 minutes to arrive at the scene of an emergency?

b. What is the probability that the fire department takes between 4 minutes and 7 minutes to arrive at the scene of an emergency?

29. Boxes of cereal are filled by a maching. Tests of the machine's accuracy show that the amount of cereal in each box varies. The weights are normally distributed with a mean of 20 ounces and a standard deviation of 0.25 ounce.

a. Find the z-scores for weights of 19.4 ounces and 20.4 ounces.

b. What is the probability that a randomly selected cereal box weighs between 19.4 ounces and 20.4 ounces. Explain your reasoning.