Normal Random Variables

5.1 Using the Z Table

**Continuous Probability Distributions**

In this section, we work with random variables that are continuous. A **continuous random variable** can assume any numerical value within some interval or intervals. Let X be one such continuous random variable, then:

The probability function $f(x)$ for X, a continuous random variable, is called a **probability density function (pdf).**

> **A probability density function must satisfy the following properties:**
> 1. The total area under the curve must equal 1.
>
> 2. Every point on the curve must have a vertical height that is 0 or greater. (That is, the curve cannot fall below the *x*-axis.)

The area beneath the curve $f(x)$ between two points (a) and (b) is equal to the probability that the random variable x takes on a value between (a) and (b). That is: $P(a < x < b) = A$ = the area between points (a) and (b) under the curve $f(x)$.
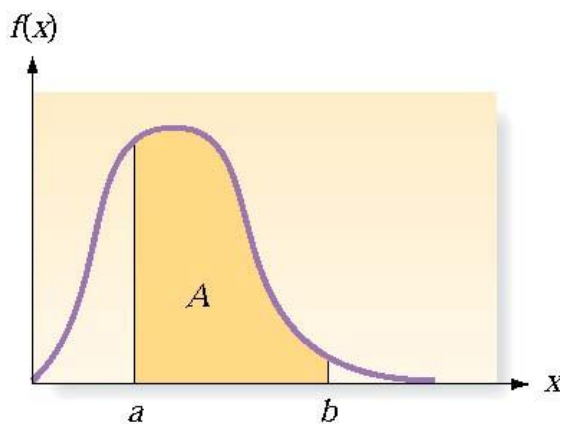
We could use calculus to find these areas; however, in this course we will use tables or software to find the areas/probabilities.

There is an important property regarding areas of this type that we should cover. Since a continuous random variable can take on an infinite, uncountable number of values there is zero probability that x will take on any specific value. In other words, P(x = a) = 0 for any *a* in the function's domain.

This means that unlike discrete distributions, P(a < x < b) = $P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b)$

Example 74 Male weights are normally distributed with a mean of 172 lbs and a standard deviation of 29 lbs. What is the probability that a randomly selected man from the population will weigh exactly 182 lbs?

Solution: P(x = 182) = 0, since weights are normally distributed and thus continuous, we will have zero probability of selecting a male who weighs exactly 182 lbs.
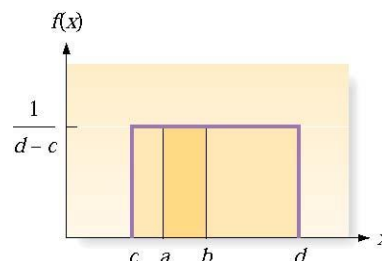
The Uniform Distribution

The uniform distribution is a continuous probability distribution with the following properties:

- *X* can take on any value between *c* and *d* with equal probability
  = *1/(d - c)*

- For two values *a* and *b*

  $$P(a < x < b) = \frac{b-a}{d-c}$$

  $$c \leq a < b \leq d$$

The Normal Distribution

One of the most common continuous distributions is the __normal distribution__. It is bell-shaped, so we call it the bell curve. It is perfectly symmetrical around its mean. Many natural phenomena can be modeled by the normal distribution.

The Normal Probability Distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \quad \text{where } -\infty < x < \infty$$

Note: $e \approx 2.7183$ and $\pi \approx 3.1416$

There exists many different random variables that follow a normal or approximately normal distribution (each combination of μ and σ produces a unique normal curve). We could create a table to give us areas from under the normal curve for each different random variable we encounter, but this would be very time consuming and thus limiting. To fix this problem we can standardize our random variables, which will allow us to use ***one normal table to solve every possible problem***:

The standardized normal random variable

$$z = \frac{x - \mu}{\sigma}$$

The mean of z is 0, and the standard deviation is 1.
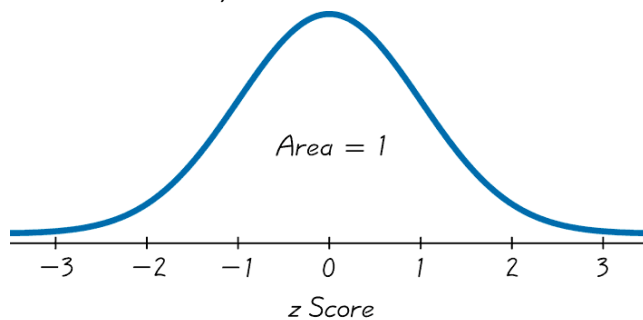
The standard normal distribution has three important properties:

1. It is bell-shaped.
2. It has a mean equal to 0.
3. It has a standard deviation equal to 1.

It is extremely important to develop the skill of finding areas (or probabilities or relative frequencies) corresponding to various regions under the graph of the standard normal distribution.

Using the Standard - Normal Probability Table:
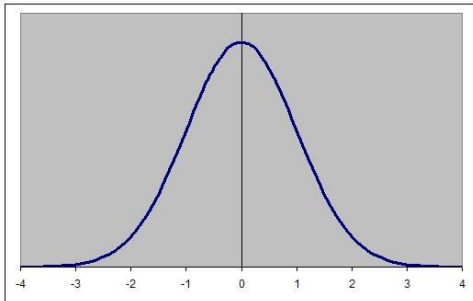
Example 75: Use the z table (standard normal probability table) to find the probability that a standard normal random variable is between 0 and 1 ($P(0 < Z < 1)$ = ?). What should it be using the empirical rule?

Example 76: Find $P(-1 < Z < 2)$

Example 77: Find $P(Z > 2.34)$

Example 78: Find $P(-2.05 < Z < -1.65)$
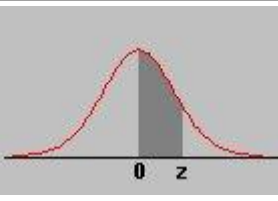
Example 79: Find $P(Z > -1.25)$

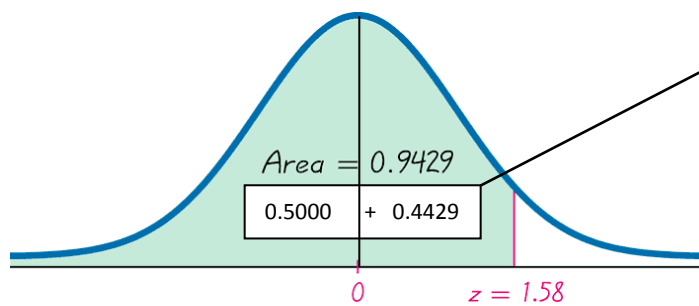Example 80: Thermometers have an average reading of 0 degrees and a standard deviation of 1 degree when inserted into freezing water. Assuming that these readings are normally distributed and one thermometer is randomly selected and tested, find the probability that, at the freezing point of water, the reading is less than 1.58 degrees.

Solution: $P(z < 1.58)$ = We look up 1.58 in the table to find the area from 1.58 to the mean.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545

Area = 0.9429

| 0.5000 | + 0.4429 |

0 $z = 1.58$

The probability that the chosen thermometer will measure freezing water as less than 1.58 degrees is 0.9429.

Note: The Empirical Rule applies to the Normal distribution, so we know that it is rare to find data points outside of ±3 standard deviations from the mean. We will consider an event to be rare if it has less than a 5% chance of occurring.

Example 81: If Z is a standard normal variable, find the probability that Z lies between 0.7 and 1.98.

    A.   0.2175
    B.   -0.2181
    C.   1.7341
    D.   0.2181

Solution: D. 0.2181

Example 82: Thermometers have an average reading of 0 degrees and a standard deviation of 1 degree when inserted into freezing water. Assuming that these readings are normally distributed and one thermometer is randomly selected and tested, find the probability that it reads (at the freezing point of water) above –1.23 degrees.
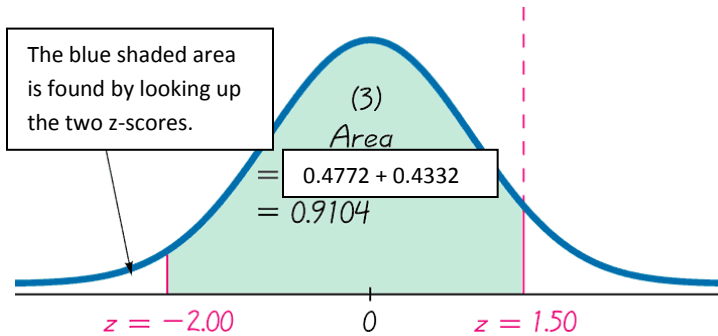
Solution: **P (z > –1.23) = 0.8907, 89.07% of the thermometers have readings above –1.23 degrees.**

	0.00	0.01	0.02	0.03
1.1	0.3643	0.3665	0.3686	0.3708
1.2	0.3849	0.3869	0.3888	0.3907
1.3	0.4032	0.4049	0.4066	0.4082

The right side of the curve is 0.5000

$0.3907 + 0.5000$

$= 0.8907$

$z = -1.23$ 0
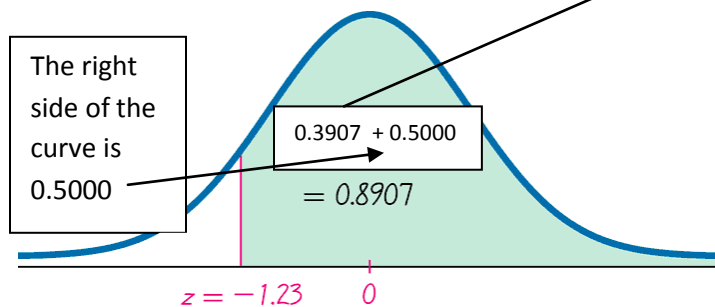
Example 83: A thermometer is randomly selected. Find the probability that it reads (at the freezing point of water) between –2.00 and 1.50 degrees.

Solution:

The blue shaded area is found by looking up the two z-scores.

(3)
Area
$=$ 0.4772 + 0.4332
$= 0.9104$

$z = -2.00$ 0 $z = 1.50$

## 5.2 Probability Using the Normal Distribution

Using the Standard – Normal probability table for non-standard normal random variables:

This section presents methods for working with normal distributions that are not standard. That is, the mean is not 0 or the standard deviation is not 1, or both. The key concept is that we can use the z-score formula to convert non-standard normal random variables into standard normal random variables. This will allow us to use the methods of the previous section to find probabilities.

Steps to finding Probability corresponding to a Normal Random Variable:

1. Sketch the Bell Curve and label its mean and standard deviation.
2. Shade the area you are looking to find.
3. Convert the boundaries (x values) of the shaded area into z values using $z = \dfrac{x - \mu}{\sigma}$
4. Use the Z – table to find the desired area.

Example 84:   The distance a toy car travels on a single charge is normally distributed with an average of 3,000 yards and a standard deviation of 50 yards (i.e., μ = 3,000 and σ = 50). What is the probability that the car will go more than 3,100 yards without recharging?
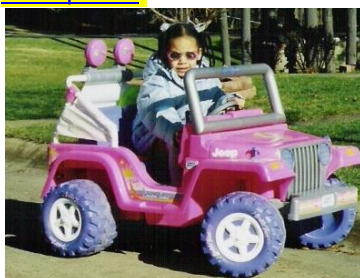
Example 85: Lengths of pregnancies are normally distributed with a mean of 268 days and a standard deviation of 15 days; find the probability of a woman carrying for a term of at least 308 days.

Example 86: Student loan debt for FIU graduates is normally distributed. In 2010, the average debt for those borrowing to pay for school was $16,026 with a standard deviation of $8,623. Find the probability that a graduate from the class of 2010 with student loans has less than $25,000 in debt.


Example 86.5: Starting salaries for business majors in the US are normally distributed with a mean of $44,000 and a standard deviation of $5,300. Find the probability that a randomly selected business school graduate earns less than $30,000 or more than $60,000.

5.3 Finding Percentiles of the Normal Curve (Using the Table in Reverse)

Using the normal table in reverse:

There are times when we are asked to determine the value that corresponds to a given area or amount of probability. For example, if you want to know what height a woman has to be to be taller than 90% of the population of all women. We know the probability or area under the curve, but we do not know what x value corresponds to that area. To solve this sort of problem we need to know the mean and standard deviation for the random variable x (in this example x = the height of women).

$Area =$ 0.90

0

$z = ?$
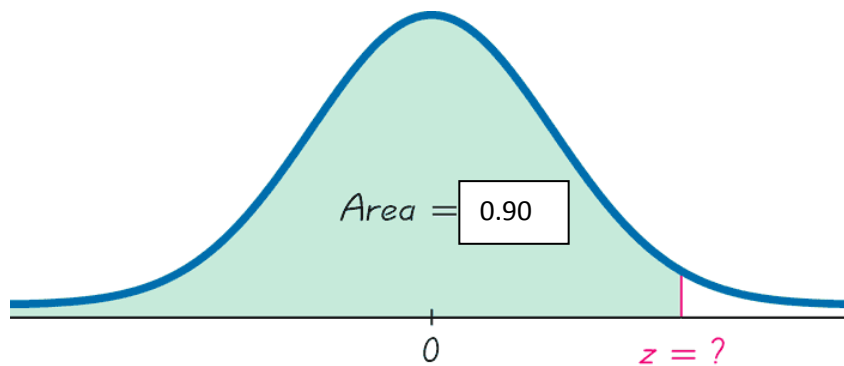
Example 87: Human heights are normally distributed. Assume the average height for women is 63.6 inches with a standard deviation of 2.5 inches; find the height that separates the bottom 90% of all women from the top 10%.

Example 88: The design of caskets depends on average human heights. If men's heights are normally distributed, and have an average of 69 inches with a standard deviation of 2.8 inches, find the casket length that will fit only the shortest 25% of men.

Example 88.5: A recent study of the influence of diet and exercise on skeletal muscle and visceral adipose tissue in obese men found that men who used diet and resistance training for 16 weeks changed their waist circumference by an average of -4.69 inches. The standard deviation was 1.57 inches. The changes in waist circumferences were normally distributed. Find the change in waist circumference that corresponds to the 5th percentile (the change in waist circumference that corresponds to the smallest 5% change).

Example 89: Find the cutoff z-scores for the bottom 2.5% and the top 2.5% of the standard normal curve.
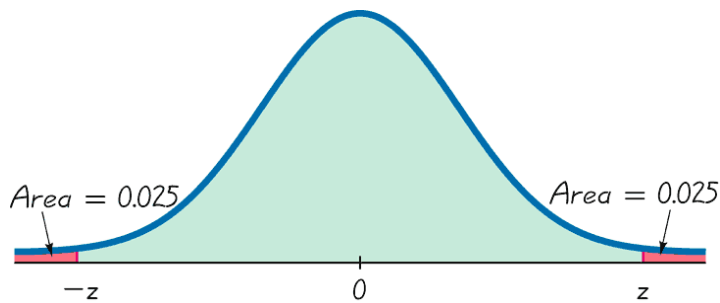
Solution:

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389

1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817

.5 - 0.025 = 0.4750

Area = 0.025   0.4750   0.4750   Area = 0.025

−1.96   0   1.96
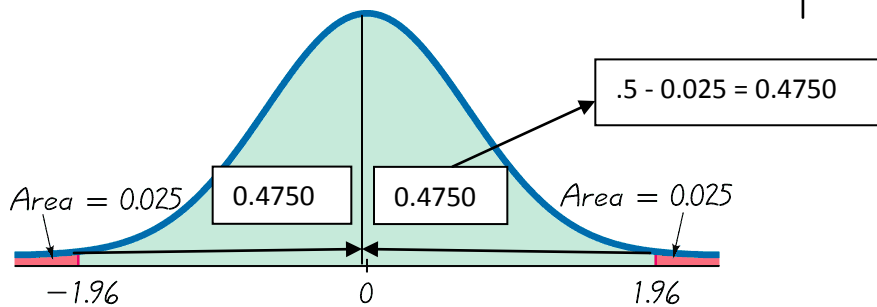
Example 90: The safe load for a water taxi was found to be 3500 pounds. The mean weight of a passenger is assumed to be 140 pounds. Assume the worst case that all passengers are men. Assume also that the weights of the men are normally distributed with a mean of 172 pounds and standard deviation of 29 pounds. If one man is randomly selected, what is the probability he weighs less than 174 pounds?
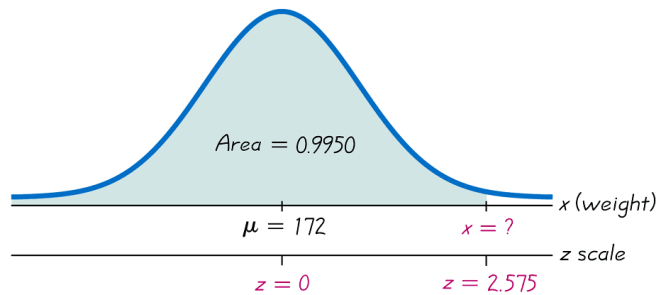
Solution: $z = \dfrac{174 - 172}{29} = 0.07$ **P (x < 174 lb.) = P(z < 0.07)**

μ = 172
σ = 29

Area = 0.5279
$\mu = 172$ $x = 174$ — x (weight)
— z scale
$z = 0$ $z = 0.07$
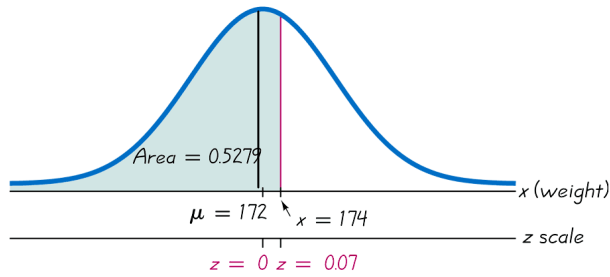
Example 91: Use the male weight data from the previous example to determine what weight separates the lightest 99.5% from the heaviest 0.5%?

Solution:

Area = 0.9950
$\mu = 172$ $x = ?$ — x (weight)
— z scale
$z = 0$ $z = 2.575$

$x = \mu + (z \bullet \sigma)$

$x = 172 + (2.575 \bullet 29)$

$x = 246.675$ (247 rounded)

**The weight of 247 pounds separates the lightest 99.5% from the heaviest 0.5%**

Things to keep in mind:

1. Don't confuse z scores and areas. z scores are distances along the horizontal scale, but areas are regions under the normal curve. The z-table lists z scores in the left column and across the top row, but areas are found in the body of the table.
2. Choose the correct (right/left) side of the graph when finding percentiles.
3. A z-score must be negative whenever it is located on the left half of the normal distribution.
4. Areas (or probabilities) are positive or zero values, but they are never negative.

5.4 Normal as Approximation to Binomial

When working with binomial probability the calculations can be very labor intensive. For example, try to find the probability that when guessing on a 50 question multiple choice exam, that you miss at most 20 questions. Tables and software can be used to speed up the work involved in problems dealing with many cumulative calculations; however, most tables do not have n values greater than 25. In cases where the sample size is too large to use the binomial table, computers can be used to perform the calculations, but they are not always readily available. This creates the need for an approximation technique.

The normal approximation method involves using a bell curve with a mean of $\mu = n(p)$ and standard deviation of $\sigma = \sqrt{npq}$ in place of the exact binomial distribution the problem involves. From the images below, you can see that when the p for the binomial random variable is 0.5 (the image on the right side) the approximation is really quite good. However, it is not as good when p = 0.10 (the image on the left side).

Graph A: B (20, 0.1)

Graph B: B (20, 0.5)
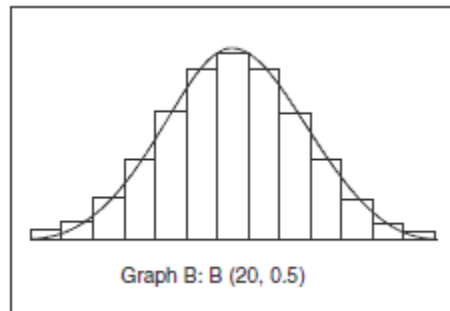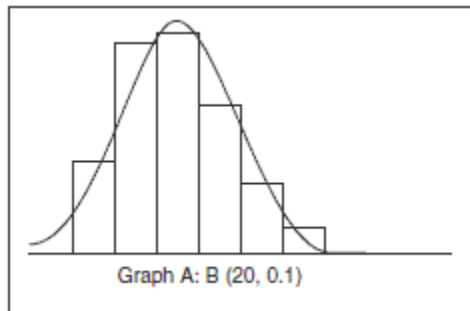
The further p is away from 0.5, the less well the approximation performs. Once p gets too far from 0.5, it is best not to use the approximation any longer. As a precaution, it is recommended that we check to see if $n(p) \geq 15$ and $n(q) \geq 15$. As long as those two conditions are met, the approximation should be reasonably good.

The approximation can be improved by using a technique called **Continuity Correction**. This technique involves adding or subtracting 0.5 to your x-value in order to correct for the difference between discrete distributions (like the binomial distribution) and continuous distributions (like the normal distribution).

The issue has to do with the fact that in discrete distributions, the probability rectangle for an X value like x = 9 actually extends from 8.5 until 9.5 on the x-axis. This is not the case for continuous random variables. When dealing with probabilities like P(9 < x < 11) for continuous distributions, X = 9 and x = 11 are infinitesimally small points on the number line which bound our probability area. To compensate for this difference, we will calculate the probability x is between 9 and 11 by actually calculating the

probability that x is between 8.5 and 11.5 when using the normal approximation to find binomial probabilities like P(9 < X < 11).
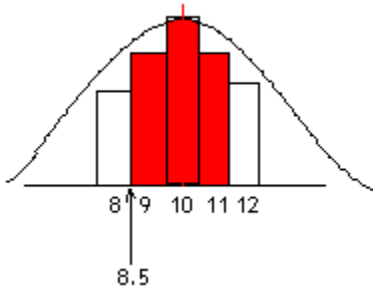
Example 91.5: According to a poll conducted in 2012, 85% of recent college graduates are planning to move back into their parent's home after graduation. Use the results of the survey to find the probability, that among 50 randomly selected recent graduates, at least 40 of them move back in with their parents after graduation.