Normal Approximations to Binomial Distributions

Normal Approximation

The normal distribution is used to approximate the binomial distribution when it would be impractical to use the binomial distribution to find a probability.

Normal Approximation to a Binomial Distribution

If $np \geq 5$ and $nq \geq 5$, then the binomial random variable x is approximately normally distributed with mean

$$\mu = np$$

and standard deviation

$$\sigma = \sqrt{npq}.$$

Normal Approximation

Example:

Decided whether the normal distribution to approximate x may be used in the following examples.

1. Thirty-six percent of people in the United States own a dog. You randomly select 25 people in the United States and ask them if they own a dog.

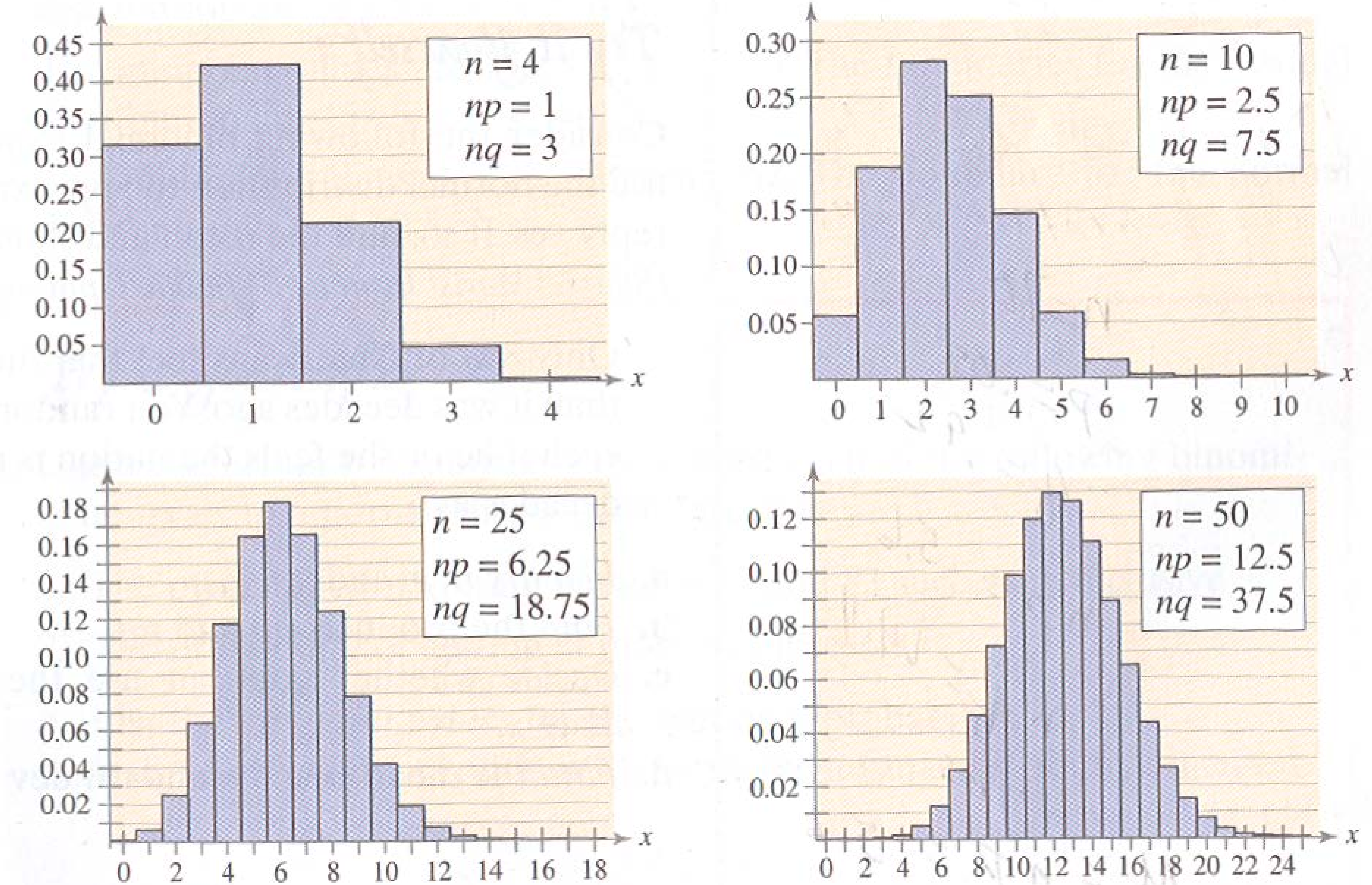
$$np = (25)(0.36) = 9$$
$$nq = (25)(0.64) = 16$$

Because np and nq are greater than 5, the normal distribution may be used.

2. Fourteen percent of people in the United States own a cat. You randomly select 20 people in the United States and ask them if they own a cat.

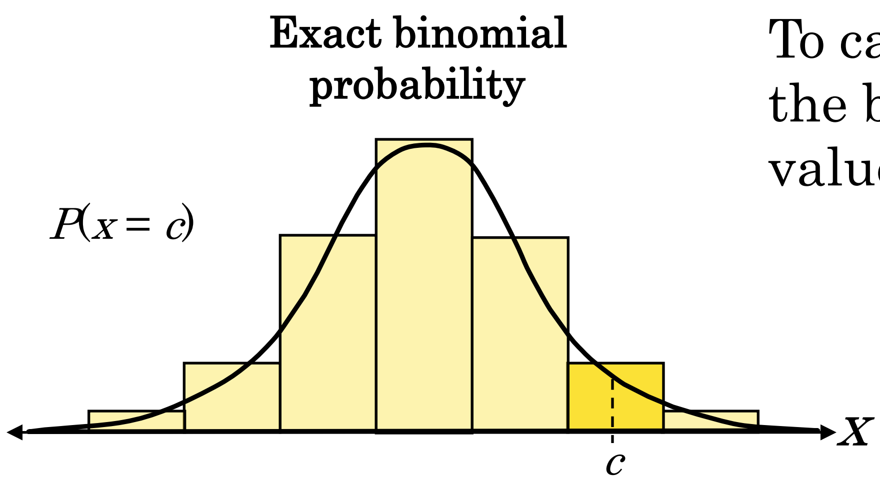
$$np = (20)(0.14) = 2.8$$
$$nq = (20)(0.86) = 17.2$$

Because np is not greater than 5, the normal distribution may NOT be used.

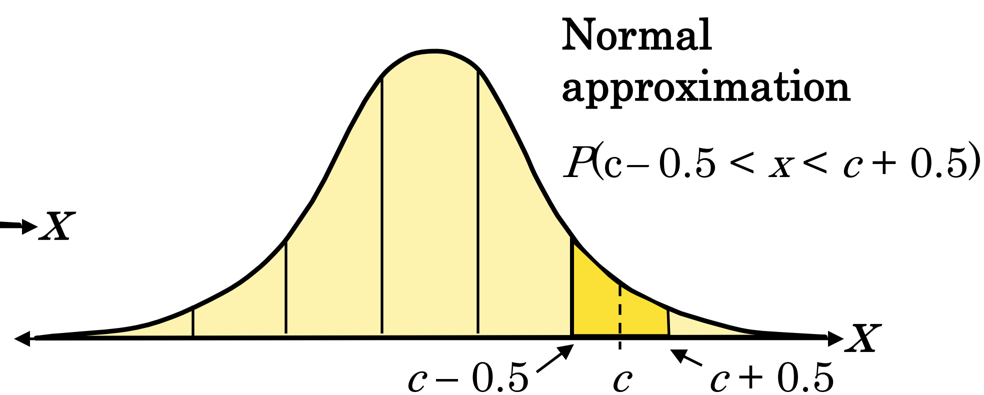

n = 4
np = 1
nq = 3
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0 1 2 3 4
x
n = 10
np = 2.5
nq = 7.5
0.30
0.25
0.20
0.15
0.10
0.05
0 1 2 3 4 5 6 7 8 9 10
x
n = 25
np = 6.25
nq = 18.75
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0 2 4 6 8 10 12 14 16 18
x
n = 50
np = 12.5
nq = 37.5
0.12
0.10
0.08
0.06
0.04
0.02
0 2 4 6 8 10 12 14 16 18 20 22 24
x

Correction for Continuity

The binomial distribution is discrete and can be represented by a probability histogram.



To calculate *exact* binomial probabilities, the binomial formula is used for each value of x and the results are added.



When using the *continuous* normal distribution to approximate a binomial distribution, move 0.5 unit to the left and right of the midpoint to include all possible x-values in the interval.

This is called the **correction for continuity**.

Correction for Continuity

Example:
Use a correction for continuity to convert the binomial intervals to a normal distribution interval.

1. The probability of getting between 125 and 145 successes, inclusive.

   The discrete midpoint values are 125, 126, …, 145.
   The continuous interval is $124.5 < x < 145.5$.

2. The probability of getting exactly 100 successes.

   The discrete midpoint value is 100.
   The continuous interval is $99.5 < x < 100.5$.

3. The probability of getting at least 67 successes.

   The discrete midpoint values are 67, 68, ….
   The continuous interval is $x > 66.5$.

Guidelines

Using the Normal Distribution to Approximate Binomial Probabilities

In Words	***In Symbols***
1. Verify that the binomial distribution applies.	Specify n, p, and q.
2. Determine if you can use the normal distribution to approximate x, the binomial variable.	Is $np \geq 5$? Is $nq \geq 5$?
3. Find the mean μ and standard deviation σ for the distribution.	$\mu = np$ $\sigma = \sqrt{npq}$
4. Apply the appropriate continuity correction. Shade the corresponding area under the normal curve.	Add or subtract 0.5 from endpoints.
5. Find the corresponding z-value(s).	$z = \dfrac{x - \mu}{\sigma}$
6. Find the probability.	Use the Standard Normal Table.

Approximating a Binomial Probability

Example:

Thirty-one percent of the seniors in a certain high school plan to attend college. If 50 students are randomly selected, find the probability that less than 14 students plan to attend college.

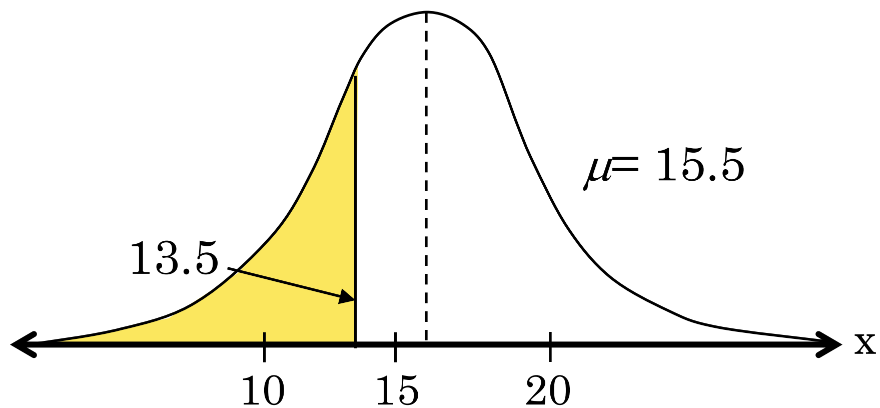
$$\left.\begin{aligned} np &= (50)(0.31) = 15.5 \\ nq &= (50)(0.69) = 34.5 \end{aligned}\right\}$$

The variable x is approximately normally distributed with $\mu = np = 15.5$ and $\sigma = \sqrt{npq} = \sqrt{(50)(0.31)(0.69)} = 3.27$.

$$P(x < 13.5) = P(z < -0.61)$$
$$= 0.2709$$

Correction for continuity

$$z = \frac{x - \mu}{\sigma} = \frac{13.5 - 15.5}{3.27} = -0.61$$



The probability that less than 14 plan to attend college is 0.2709.

Approximating a Binomial Probability

Example:

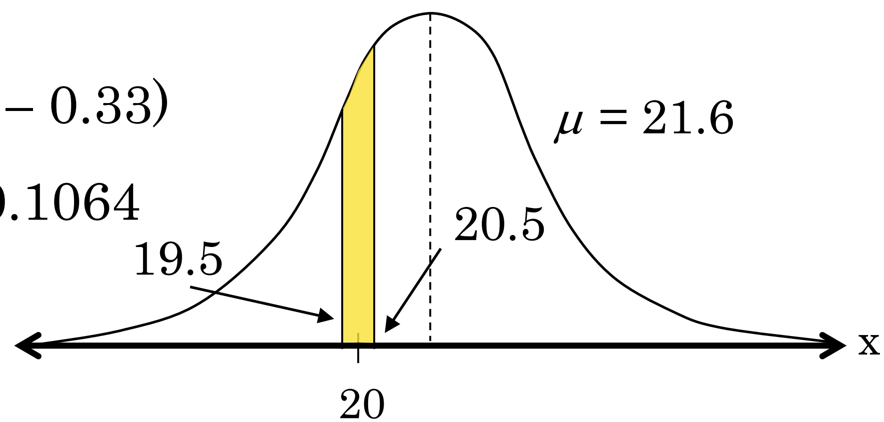
A survey reports that forty-eight percent of US citizens own computers. 45 citizens are randomly selected and asked whether he or she owns a computer. What is the probability that exactly 20 say yes?

$$\left.\begin{aligned} np &= (45)(0.48) = 21.6 \\ nq &= (45)(0.52) = 23.4 \end{aligned}\right\} \quad \begin{aligned} \mu &= 21.6 \\ \sigma &= \sqrt{npq} = \sqrt{(45)(0.48)(0.52)} = 3.35 \end{aligned}$$

$$P(19.5 < x < 20.5) = P\ (-0.63 < z - 0.33)$$

$$= 0.3707 - 0.2643 = 0.1064$$

Correction for continuity

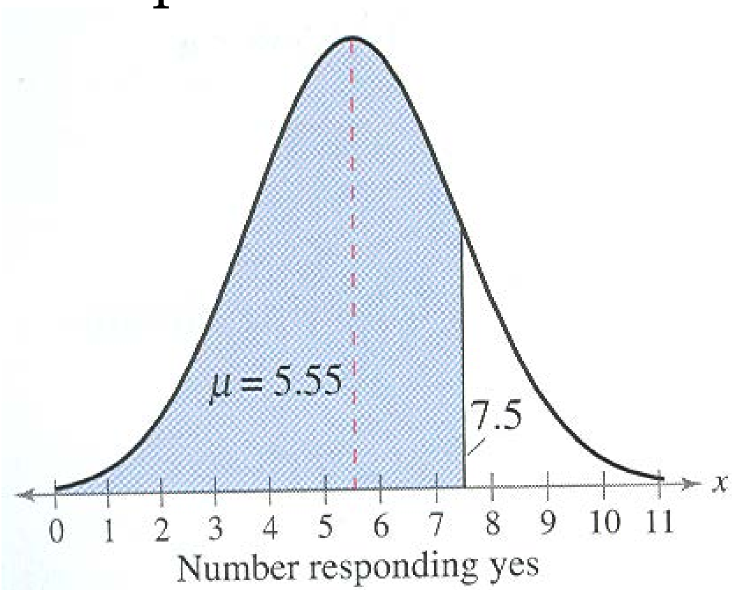


The probability that exactly 20 US citizens own a computer is 0.1064.

Example:

Thirty-seven percent of Americans say they always fly an American flag on the Fourth of July. You randomly select 15 Americans and ask each if he or she flies an American flag on the Fourth of July. What is the probability that fewer than eight of them reply yes?

You can use a normal distribution with $\mu = 5.55$ and $\sigma \approx 1.87$ to approximate the binomial distribution. By applying the continuity correction, you can rewrite the discrete probability $P(x < 8)$ as $P(x < 7.5)$. The graph on the next slide shows a normal curve with $\mu = 5.55$ and $\sigma \approx 1.87$ and a shaded area to the left of 7.5. The z-score that corresponds to x = 7.5 is



$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{7.5 - 5.55}{1.87} \approx 1.04$$

Using the Standard Normal Table,

$P(z<1.04) = 0.8508$

So, the probability that fewer than eight people respond yes is 0.8508

Example:

Twenty-nine percent of Americans say they are confident that passenger trips to the moon will occur during their lifetime. You randomly select 200 Americans and ask if he or she thinks passenger trips to the moon will occur in his or her lifetime. What is the probability that at least 50 will say yes?

SOLUTION: Because $np = (200)(0.29) = 58$ and

$nq = (200)(0.71) = 142$, the binomial variable x is approximately normally distributed with

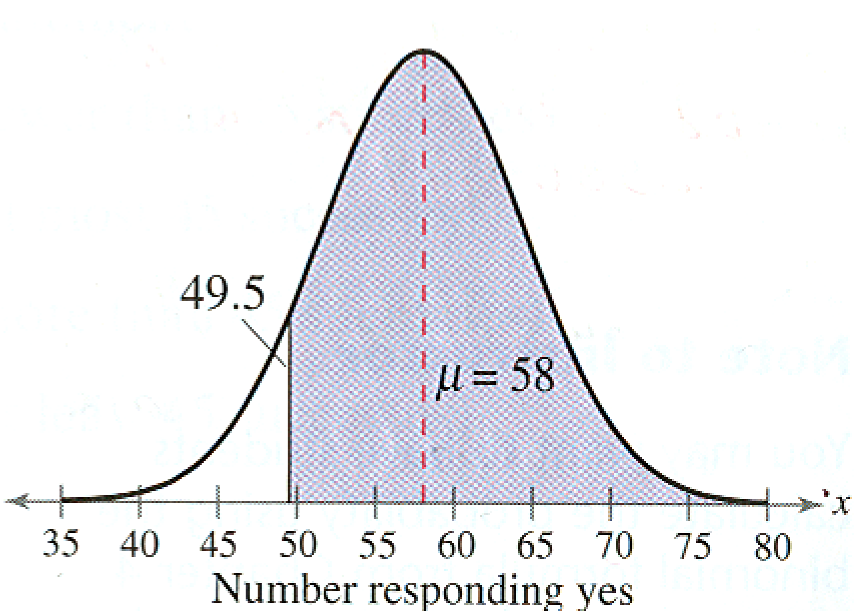
$$\mu = np = 58$$

and

$$\sigma = \sqrt{npq} = \sqrt{200 \cdot 0.29 \cdot 0.71} \approx 6.42$$

Using the correction for continuity, you can rewrite the discrete probability P $(x \geq 50)$ as the continuous probability P $(x \geq 49.5)$. The graph shows a normal curve with $\mu = 58$ and $\sigma = 6.42$, and a shaded area to the right of 49.5.

The z-score that corresponds to 49.5 is



$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{49.5 - 58}{6.42} \approx -1.32$$

So, the probability that at least 50 will say yes is:

$P(x \geq 49.5) = 1 – P(z \leq -1.32)$
$= 1 – 0.0934 = 0.9066$

Study Tip

In a discrete distribution, there is a difference between $P\ (x \geq c)$ and $P(\ x > c)$. This is true because the probability that x is exactly c is not zero. IN a continuous distribution, however, there is no difference between $P\ (x \geq c)$ and $P\ (x > c)$ because the probability that x is exactly c is zero.

Example:

A survey reports that 48% of Internet users use Netscape as their browser. You randomly select 125 Internet users and ask each whether he or she uses Netscape as his or her browser. What is the probability that exactly 63 will say yes?

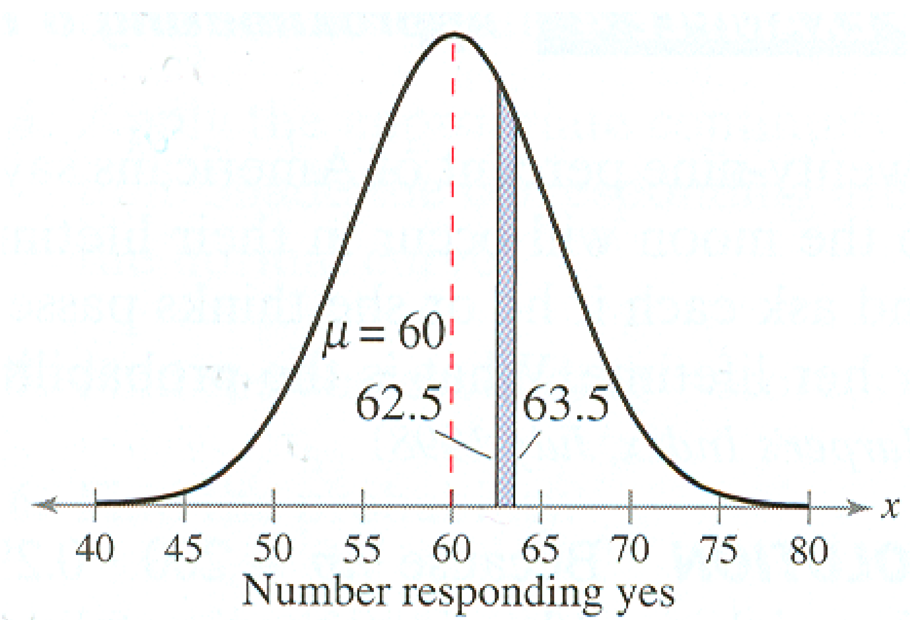
SOLUTION: Because $np = (125)(0.48) = 60$ and $nq = (125)(0.52) = 65$, the binomial variable x is approximately normally distributed with

$\mu = np = 60$ and

$$\sigma = \sqrt{npq} = \sqrt{125 \cdot 0.48 \cdot 0.52} \approx 5.59$$

Using the correction for continuity, you can rewrite the discrete probability P ($x \geq 63$) as the continuous probability P ($62.5 < x < 63.5$). The graph shows a normal curve with $\mu = 60$ and $\sigma = 5.59$, and a shaded area between 62.5 and 63.5.

The z-scores that corresponds to 62.5 and 63.5



$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{62.5 - 60}{5.59} \approx 0.45$$

$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{63.5 - 60}{5.59} \approx 0.63$$

So, the probability that at least 50 will say yes is:

$$P(62.5 < x < 63.5) = P\,(0.45 < z < 0.63)$$

$$= P(z < 0.63) - P(z < 0.45)$$

$$= 0.7357 - 0.6736$$

$$= 0.0621$$

There is a probability of about 0.06 that exactly 63 of the Internet users will say they use Netscape.

Example:

Suppose that we roll 2 dice 180 times. Let E be the event that we roll two fives no more than once.

(a) Find the exact probability of E.

(b) Approximate $\mathbb{P}(E)$ using the normal distribution.
(c) Approximate $\mathbb{P}(E)$ using the Poisson distribution.

Example:

About 10% of the population is left-handed. Use the normal distribution to approximate the probability that in a class of 150 students,

(a) at least 25 of them are left-handed.
(b) between 15 and 20 are left-handed.

Example:

A teacher purchases a box with 50 markers of colors selected at random. The probability that marker is black is 0.6, independent of all other markers. Knowing that the probability of there being more than N black markers is greater than 0.2 and the probability of there being more than N + 1 black markers is less than 0.2, use the normal approximation to calculate N.