Lecture Three

Section 3.1 – Distribution of the Sample Mean / Proportion

Statistics such as $\overline{x}$ are random variables since their value varies from sample to sample. As such, they have probability distributions associated with them. We focus on the shape, center and spread of statistics such as $\overline{x}$.

Definitions

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.

Properties

- Sample means target the value of the population mean. (That is, the mean of the sample means is the population mean. The expected value of the sample mean is equal to the population mean.)
- The distribution of the sample means tends to be a normal distribution.

Sampling Distributions

Step 1: Obtain a simple random sample of size *n*.
Step 2: Compute the sample mean.
Step 3: Assuming we are sampling from a finite population, repeat Steps 1 and 2 until all simple random samples of size *n* have been obtained.

Definition

The ***sampling distribution of the variance*** is the distribution of sample variances, with all samples having the same sample size *n* taken from the same population. (The sampling distribution of the variance is typically represented as a probability distribution in the format of a table, probability histogram, or formula.)

Biased Estimators

Sample *medians*, *ranges* and *standard deviations* are biased estimators. That is they do NOT target the population parameter.

Note: the bias with the standard deviation is relatively small in large samples so *s* is often used to estimate.

Unbiased Estimators

Sample means, variances and proportions are ***unbiased estimators***. That is they target the population parameter.
These statistics are better in estimating the population parameter

- *Mean* $\overline{x}$
- *Variance* s^2
- *Proportion* $\hat{p}$

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

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

Notation

The mean of the sample means	$\mu_{\overline{x}} = \mu$
The standard deviation of sample mean	$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$

The Shape of the Sampling Distribution of $\overline{x}$ If *X* is Normal

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

Applying

Individual value: When working with individual value from a normally distributed population, use

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

Sample value: When working with a mean for some sample (or group), use the value of $\frac{\sigma}{\sqrt{n}}$ for the standard deviation of the sample means. Use $z = \dfrac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

Example

Assume the population of weights of men is normally distributed with a mean of 172 lb. and a standard deviation of 29 lb.

a) Find the probability that if an *individual* man is randomly selected, his weight is greater than 175 lb.

b) Find the probability that *20 randomly selected men* will have a mean weight that is greater than 175 lb. (so that their total weight exceeds the safe capacity of 3500 pounds).

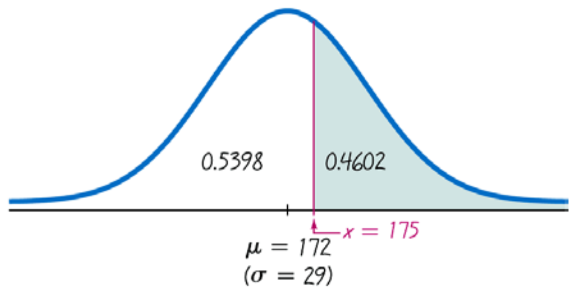
Solution

a) We are dealing with an individual value from a normally distributed population.

$$z = \frac{x-\mu}{\sigma} = \frac{175-172}{29} = 0.10$$

Using the Table (*Normal Distribution*), $A_1 = 0.5398$

Therefore, the shaded region area is: $A = 1 - 0.5398 \underline{= 0.4602}$



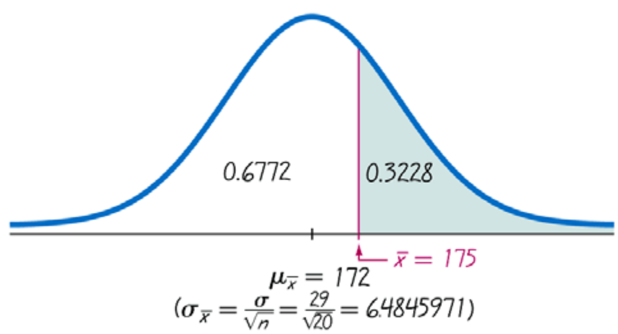
The probability of a randomly selected man weighing more than 175 lb. is 0.4602. (Calculator result is 0.4588)

b) Although the sample size is not greater than 30. We use a normal distribution because the original population of men has a normal distribution.

$$\mu_{\overline{x}} = \mu = 172$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{29}{\sqrt{20}} = 6.4845971$$

$$z = \frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{175-172}{6.4845971} = 0.46$$



$$A_2(<0.46) = 0.6772$$

Therefore, the shaded region area is: $A = 1 - 0.6772 \underline{= 0.3228}$

So, the probability that the 20 men have a mean weight greater than 175 lb. is 0.3228. (Calculator result is 0.3218)

✓ There is 0.4602 probability that an individual man will weigh more than 175 lb., and there is a 0.3228 probability that 20 men will have a mean weight of more than 175 lb. Given that the safe capacity of the water taxi is 3500 pounds, there is a fairly good chance (with probability 0.3228) that it will be overloaded with 20 randomly selected men. The capacity of 20 passengers is just not safe enough.

Example

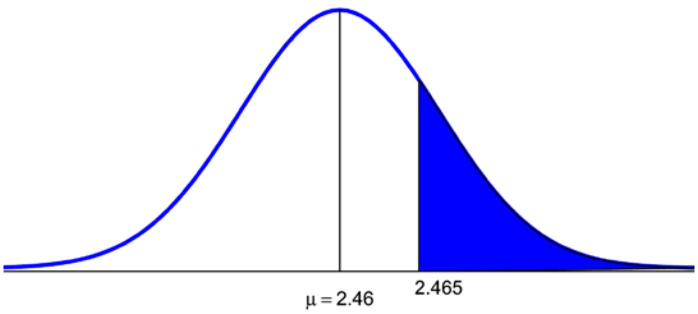
The weights of pennies minted after 1982 are approximately normally distributed with mean 2.46 grams and standard deviation 0.02 grams.
What is the probability that in a simple random sample of 10 pennies minted after 1982, we obtain a sample mean of at least 2.465 grams?

Solution

$$\mu_{\overline{x}} = 2.46, \quad \sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.02}{\sqrt{10}} = .0063$$

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.465 - 2.46}{0.0063} = 0.79$$

$$P(z > 0.79) = 1 - 0.7852 \underline{= 0.2148}$$



Example

The following table and histogram give the probability distribution for rolling a fair die.

Solution

$$\underline{\mu = 3.5}$$

$$\underline{\sigma = 1.708}$$

The population distribution is ***NOT*** normal.

Face on Die	*Relative Frequency*
1	0.1667
2	0.1667
3	0.1667
4	0.1667
5	0.1667
6	0.1667

Correction for a Finite Population

When sampling without replacement and the sample size n is greater than 5% of the finite population of size N (that is, $n > 0.05N$), adjust the standard deviation of sample means by multiplying it by the *finite population correction factor*:

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} \cdot \underbrace{\sqrt{\frac{N-n}{N-1}}}_{\textit{finite population correction factor}}$$

Example

Cans of regular Coke are labeled to indicate that they contain 12 *oz*. The corresponding sample statistics are n = 36 and $\overline{x} = 12.19\ oz$. If the Coke cans filled so that $\mu = 12\ oz$ and the population standard deviation is $\sigma = 0.11\ oz$ (based on the sample result), find the probability that a sample of 36 cans will have a mean of 12.19 *oz*. or greater. Do these results suggest that the Coke cans are filled with an amount greater than 12.00 *oz*.?

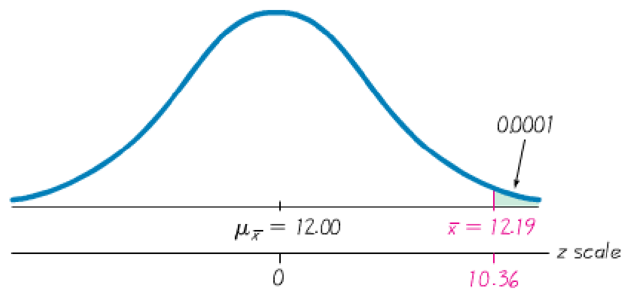
Solution

Because the sample size n = 36 exceeds 30, we apply the central limit theorem and conclude that the distribution of sample means is approximately a normal distribution with these parameters:

$$\mu_{\overline{x}} = \mu = 12$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.11}{\sqrt{36}} = 0.018333$$

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{12.19 - 12}{0.018333} = 10.36$$



Since the value of $z = 10.36$ is off the chart.
However, for values of z above 3.49, we use 0.9999 for the cumulative left area.
Therefore, we conclude that the shaded region is 0.0001.

✓ The result shows that there is an extremely small probability of getting a sample mean of 12.19 oz. or greater when 36 cans are randomly selected. It appears that the company has found a way to ensure that very few cans have less than 12 oz. while not wasting very much of their product.

Example

It is not totally unreasonable to think that screws labeled as being 3/4 in. in length would have a mean length that is somewhat close to 3/4 in. The lengths of a sample of 50 such screws have a mean length of 0.7468 in. Assume that the population of all such screws has a standard deviation described by $\sigma = 0.0123\ in$.

a) Assuming that the screws have mean length of 0.75 in. as labeled, find the probability that a sample of 50 screws has a mean length of 0.7468 in. or less.

b) The probability of getting a sample mean that is "at least as extreme as the given sample mean" is twice the probability found in part (*a*). Find this probability. (Note that the sample mean of 0.7468 in. misses the labeled mean of 0.75 in. by 0.0032 in., so any other mean is at least as extreme as the sample mean if it is below 0.75 in. by 0.0032 in. or more, or if it is above 0.75 in. by 0.0032 in. or more.)

c) Based on the result in part (b), does it appear that the sample mean misses the labeled mean of 0.75 in. by a significant amount? Explain.

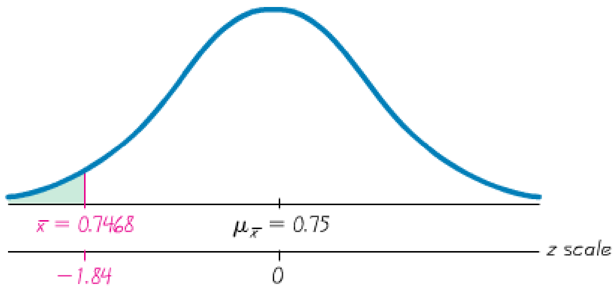
Solution

a) Because the sample size n = 50 exceeds 30, we apply the central limit theorem and conclude that the distribution of sample means is a normal distribution with these parameters:

$$\mu_{\bar{x}} = \mu = 0.75$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.0123}{\sqrt{50}} = 0.001739$$

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{0.7468-0.75}{0.001739} = -1.84$$



$$A(<-1.84) \ \underline{= 0.0329}$$

The probability of getting a sample mean of 0.7468 in. or less is 0.0329.

b) $P(\textit{at least as extreme as the given sample mean}) = 2P(\textit{part } a)$

$$= 2 \times 0.0329$$

$$\underline{= 0.0658}$$

c) The result from part (b) shows that there is a 0.0658 probability of getting a sample mean that is at least as extreme as the given sample. Using a 0.05 cutoff probability of 0.0658 exceeds 0.05, so the sample mean is not unusual. We conclude that the given sample mean does not miss the labeled mean of 0.75 in. by a substantial amount. The labeling of 0.75 in. appears to be justified.

Central Limit Theorem

Regardless of the shape of the underlying population, the sampling distribution of $\bar{x}$ becomes approximately normal as the sample size, *n*, increases.

The mean of the sampling distribution is equal to the mean of the parent population and the standard deviation of the sampling distribution of the sample mean is $\frac{\sigma}{\sqrt{n}}$ regardless of the sample size.

Example

Suppose that the mean time for an oil change at a "10-minute oil change joint" is 11.4 minutes with a standard deviation of 3.2 minutes.

a) If a random sample of $n = 35$ oil changes is selected, describe the sampling distribution of the sample mean.

b) If a random sample of $n = 35$ oil changes is selected, what is the probability the mean oil change time is less than 11 minutes?

Solution

a) $\mu_{\bar{x}} = 11.4$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3.2}{\sqrt{35}} = 0.5409$$

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{0.7468-0.75}{0.001739} \ \underline{= -1.84}$$

Point Estimate of a Population Proportion

Suppose that a random sample of size *n* is obtained from a population in which each individual either does or does not have a certain characteristic. The sample proportion, denoted $\hat{p}$ (read "*p*-hat") is given by

$$\hat{p} = \frac{x}{n}$$

where *x* is the number of individuals in the sample with the specified characteristic. The sample proportion $\hat{p}$ is a statistic that estimates the population proportion, *p*.

Example

In a Quinnipiac University Poll conducted in May of 2008, 1745 registered voters nationwide were asked whether they approved of the way Bush is handling the economy. 349 responded "yes". Obtain a point estimate for the proportion of registered voters who approve of the way Bush is handling the economy.

Solution

$$\hat{p} = \frac{x}{n} = \frac{349}{1745} \underline{= 0.2}$$

Sampling Distribution of $\hat{p}$

For a simple random sample of size *n* with population proportion *p*:

- ✓ The shape of the sampling distribution of $\hat{p}$ is approximately normal provided $np(1-p) \geq 10$.
- ✓ The mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p$
- ✓ The standard deviation of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
- ✓ The model on the previous slide requires that the sampled values are independent. When sampling from finite populations, this assumption is verified by checking that the sample size *n* is no more than 5% of the population size *N* ($n \leq 0.05N$).
- ✓ Regardless of whether $np(1-p) \geq 10$ or not, the mean of the sampling distribution of $\hat{p}$ is *p*, and the standard deviation is $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

Example

According to a *Time* poll conducted in June of 2008, 42% of registered voters believed that gay and lesbian couples should be allowed to marry. Suppose that we obtain a simple random sample of 50 voters and determine which voters believe that gay and lesbian couples should be allowed to marry. Describe the sampling distribution of the sample proportion for registered voters who believe that gay and lesbian couples should be allowed to marry.

Solution

The sample of n = 50 is smaller than 5% of the population size (all registered voters in the U.S.).

$$npq = 50(0.42)(0.58) \underline{= 12.8 \ge 10} \quad \& \quad \underline{\mu \approx 0.42}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.42(1-0.42)}{50}} \underline{= 0.0698}$$

Example

According to the Centers for Disease Control and Prevention, 18.8% of school-aged children, aged 6-11 years, were overweight in 2004.

a) In a random sample of 90 school-aged children, aged 6-11 years, what is the probability that at least 19% are overweight?

b) Suppose a random sample of 90 school-aged children, aged 6-11 years, results in 24 overweight children. What might you conclude?

Solution

n = 90 is less than 5% of the population size

$$npq = 90(0.188)(1-0.188) \underline{\approx 13.7 \ge 10} \qquad \& \qquad \underline{\mu = 0.188}$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.188(1-0.188)}{90}} \underline{= 0.0412}$$

a) $z = \dfrac{0.19 - 0.188}{0.0412} \underline{= 0.0485}$

$$P(z > 0.05) = 1 - 0.5199 \underline{= 0.4801}$$

b) $\hat{p} = \dfrac{24}{90} = 0.2667$

$$z = \frac{0.2667 - 0.188}{0.0412} \underline{= 1.91}$$

$$P(z > 1.91) = 1 - 0.9719 \underline{= 0.028}$$

We would only expect to see about 3 samples in 100 resulting in a sample proportion of 0.2667 or more. This is an unusual sample if the true population proportion is 0.188.

Distribution of Sample Means

Population (with mean μ and standard deviation σ)	***Distribution of Sample Means***	***Mean of the Sample Means***	***Standard Deviation of the Sample Means***
Normal	Normal (for any sample size n)	$\mu_{\overline{x}} = \mu$	$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$
Not normal with $n > 30$	Normal (approximately)	$\mu_{\overline{x}} = \mu$	$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$
Not normal with $n \le 30$	Not normal	$\mu_{\overline{x}} = \mu$	$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$

Exercise *Section* 3.1 – Distribution of the Sample Mean / Proportion

1. Assume that SAT scores are normally distributed with mean $\mu = 1518$ and standard deviation $\sigma = 325$.

a) If 1 SAT score is randomly selected, find the probability that it is less than 1500.

b) If 100 SAT scores are randomly selected, find the probability that they have a mean less than 1500.

c) If 1 SAT score is randomly selected, find the probability that it is greater than 1600.

d) If 64 SAT scores are randomly selected, find the probability that they have a mean greater than 1600.

e) If 1 SAT score is randomly selected, find the probability that it is between 1550 and 1575.

f) If 25 SAT scores are randomly selected, find the probability that they have a mean between 1550 and 1575.

2. Assume that weights of mean are normally distributed with a mean of 172 lb. and a standard deviation of 29 lb.

a) Find the probability that if an individual man is randomly selected, his weight will be greater than 180 lb.

b) Find the probability that 20 randomly selected men will have a mean weight that is greater than 180 lb.

c) If 20 mean have a mean weight greater than 180 lb., the total weight exceeds the 3500 lb. safe capacity of a particular water taxi. Based on the proceeding results, is this safety concern? Why or why not?

3. Membership requires an IQ score above 131.5. Nine candidates take IQ tests, and their summary results indicated that their mean IQ score is 133, (IQ scores are normally distributed with mean of 100 and a standard deviation of 15.)

a) If 1 person is randomly selected from the general population, find the probability of getting someone with an IQ score of at least 133.

b) If 9 people are randomly selected, find the probability that their mean IQ score is at least 133.

c) Although the summary results are available, the individual IQ test scores have been lost. Can it be concluded that all 9 candidates have IQ scores above 131.5 so that they are all eligible for membership?

4. For women aged 18-24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. Hypertension is commonly defined as a systolic blood pressure above 140.

a) If a woman between the ages of 18 and 24 is randomly selected, find the probability that her systolic blood pressure is greater than 140.

b) If 4 women in that age bracket are randomly selected, find the probability that their mean systolic blood pressure is greater than 140.

c) Given that part (*b*) involves a sample size that is not larger than 30, why can the central limit theorem be used?

d) If a physician is given a report stating that 4 women have a mean systolic, blood pressure below 140, can she conclude that none of the women have hypertension (with a blood pressure greater than 140)?

5. Engineers must consider the breadths of male heads when designing motorcycle helmets. Men have head breaths that are normally distributed with a mean of 6.0 in. and a standard deviation of 1.0 in.

a) If one male is randomly selected, find the probability that his head breadth is less than 6.2 in.

b) The Safeguard Helmet Company plans an initial production run of 100 helmets. Find the probability that 100 randomly selected men have a mean head breadth less than 6.2 in.

c) The production manager sees the result from part (b) and reasons that all helmets should be made for men with head breadths less than 6.2 in., because they would fit all but a few men. What is wrong with that reasoning?

6. Currently, quarters have weighs that are normally distributed with a mean 5,670 *g* and a standard deviation of 0.062 *g*. A vending machine is configured to accept only those quarters with weights between 5.550 *g* and 5.790 *g*.

a) If 280 different quarters are inserted into the vending machine, what is the expected number of rejected quarter?

b) If 280 different quarters are inserted into the vending machine, what is the probability that the mean falls between the limits of 5.550 *g* and 5.790 *g*?

c) If you own the vending machine, which result would concern you more? The result from part (*a*) or the result from part (*b*)? Why?

7. The annual precipitation amounts in a certain mountain range are normally distributed with a mean of 100 inches, and a standard deviation of 12 inches. What is the probability that the mean annual precipitation during 36 randomly picked years will be less than 112.8 inches?

8. The annual precipitation amounts in a certain mountain range are normally distributed with a mean of 72 inches, and a standard deviation of 14 inches. What is the probability that the mean annual precipitation during 49 randomly picked years will be less than 74.8 inches?

9. The weights of the fish in a certain lake are normally distributed with a mean of 13 *lb*. and a standard deviation of 6. If 4 fish are randomly selected, what is the probability that the mean weight will be between 10.6 and 16.6 *lb*.?

10. For women aged 18-24, systolic blood pressures (in mm Hg) are normally distributed with a mean of 114.8 and a standard deviation of 13.1. If 23 women aged 18-24 are randomly selected, find the probability that their mean systolic blood pressure is between 119 and 122.

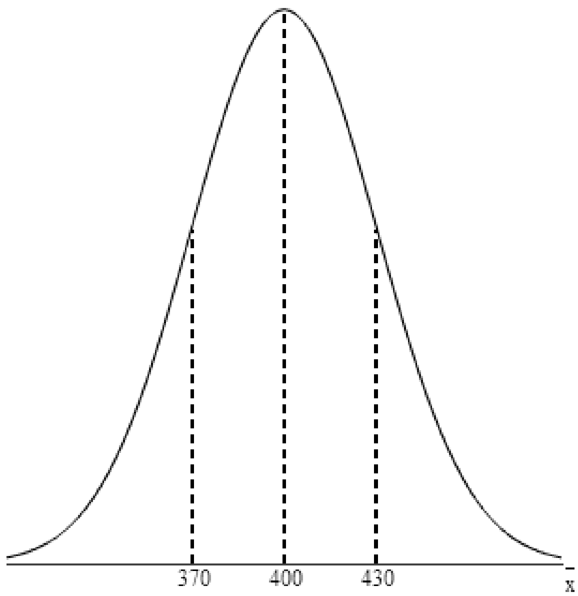
11. A study of the amount of time it takes a mechanic to rebuild the transmission for 2005 Chevy shows that the mean is 8.4 hours and the standard deviation is 1.8 hours. If 40 mechanics are randomly selected, find the probability that their mean rebuild time

a) Exceeds 8.7 hours.

b) Exceeds 8.1 hours.

12. A final exam in Math 160 has a mean of 73 with standard deviation 7.8. If 24 students are randomly selected, find the probability that the mean of their test scores is greater than 71.

13. The sampling distribution of the sample mean shown in the graph.



a) What is the value of $\mu_{\overline{x}}$?

b) What is the value of $\sigma_{\overline{x}}$?

c) If the sample size is $n = 9$, what is likely true about the shape of the population?

d) If the sample size is $n = 9$, what is the standard deviation of the population for which the sample was drawn?

14. A sample random of size $n = 81$ is obtained from a population with $\mu = 77 \quad and \quad \sigma = 18$.

a) Describe the sampling distribution of $\overline{x}$

b) What is $P(\overline{x} > 79.6)$?

c) What is $P(\overline{x} \le 72.5)$?

d) What is $P(75.1 < \overline{x} < 80.9)$?

15. The reading speed of second grade students is approximately normal, with a mean of 90 words per minute (*wpm*) and a standard deviation of 10 *wpm*.

a) What is the probability a randomly selected student will read more than 95 wpm?

b) What is the probability that a random sample of 11 second grade students results in a mean reading rate of more than 95 wpm?

c) What is the probability that a random sample of 22 second grade students results in a mean reading rate of more than 95 wpm?

d) What effect does increasing the sample size have on the probability?

Section 3.2 – Estimating a Population Proportion

In this section we present methods for using a sample proportion to estimate the value of a population proportion.

- ✓ The sample proportion is the best point estimate of the population proportion.
- ✓ We can use a sample proportion to construct a confidence interval to estimate the true value of a population proportion, and we should know how to interpret such confidence intervals.
- ✓ We should know how to find the sample size necessary to estimate a population proportion.

Definition

A ***point estimate*** is a single value (or point) used to approximate a population parameter.

For ***example***, the point estimate for the population proportion is $\hat{p} = \frac{x}{n}$, where x is the number of individuals in the sample with a specified characteristic and n is the sample size.

The ***sample proportion*** $\hat{p}$ is the best point estimate of the population proportion p.

Example

In a Pew Research Center poll, 70% of 1501 randomly selected adults in the United States believe in global warming, so the sample proportion is $\hat{p} = 0.70$. Find the best point estimate of the proportion of all adults in the United States who believe in global warming.

Solution

Because the sample proportion is the best point estimate of the population proportion, we conclude that the best point estimate of p is 0.70. When using the sample results to estimate the percentage of all adults in the United States who believe in global warming, the best estimate is 70%.

Example

In July of 2008, a Quinnipiac University Poll asked 1783 registered voters nationwide whether they favored or opposed the death penalty for persons convicted of murder. 1123 were in favor.
Obtain a point estimate for the proportion of registered voters nationwide who are in favor of the death penalty for persons convicted of murder.

Solution

Obtain a point estimate for the proportion of registered voters nationwide who are in favor of the death penalty for persons convicted of murder.

$$\hat{p} = \frac{x}{n} = \frac{1123}{1783} \underline{= 0.63}$$

Definitions

A ***confidence interval*** (or ***interval estimate***) for an unknown parameter consists of an interval of numbers based on a point estimate

A ***confidence level*** is the probability $1 - \alpha$ (often expressed as the equivalent percentage value) that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times. (The confidence level is also called ***degree of confidence***, or the ***confidence coefficient***.) is denoted $(1-\alpha)\cdot 100\%$

Most common choices are 90%, 95%, or 99%. (a = 10%), (a = 5%), (a = 1%)

For ***example***, a 95% level of confidence ($\alpha = 0.05$) implies that if 100 different confidence intervals are constructed, each based on a different sample from the same population, we will expect 95 of the intervals to contain the parameter and 5 not to include the parameter.

Confidence interval estimates for the population proportion are of the form

Point estimate ± margin of error.

The margin of error of a confidence interval estimate of a parameter is a measure of how accurate the point estimate is.

Interpreting a Confidence Interval

We must be careful to interpret confidence intervals correctly. There is a correct interpretation and many different and creative incorrect interpretations of the confidence interval $0.677 < p < 0.723$.

Correct: "We are 95% confident that the interval from 0.677 to 0.723 actually does contain the true value of the population proportion p."
This means that if we were to select many different samples of size 1501 and construct the corresponding confidence intervals, 95% of them would actually contain the value of the population proportion p.
(Note that in this correct interpretation, the level of 95% refers to the success rate of the process being used to estimate the proportion.)

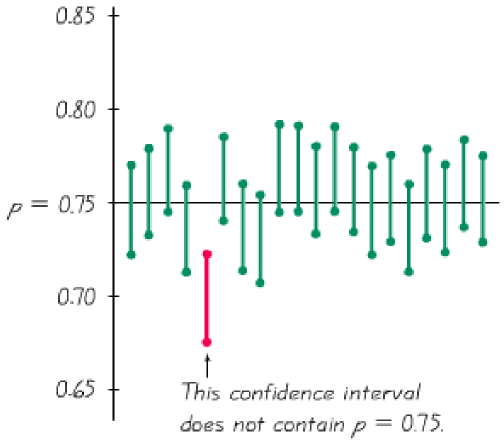
Incorrect: "There is a 95% chance that the true value of p will fall between 0.677 and 0.723." It would also be incorrect to say that 95% of sample proportions fall between 0.677 and 0.723.

The margin of error depends on ***three*** factors:

Level of confidence: As the level of confidence increases, the margin of error also increases.
Sample size: As the size of the random sample increases, the margin of error decreases.
Standard deviation of the population: The more spread there is in the population, the wider our interval will be for a given level of confidence.

Caution:

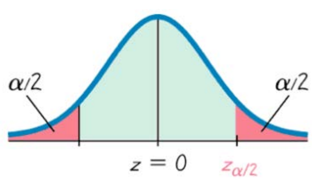
- Know the correct interpretation of a confidence interval.
- Confidence intervals can be used informally to compare different data sets, but the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of proportions.



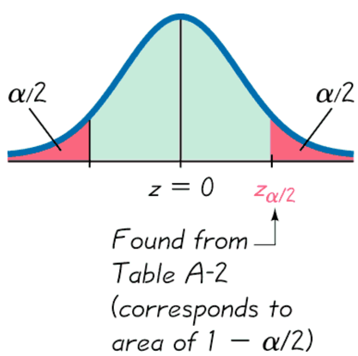
Critical Values

A standard *z* score can be used to distinguish between sample statistics that are likely to occur and those that are unlikely to occur. Such a *z* score is called a critical value. Critical values are based on the following observations:

1. Under certain conditions, the sampling distribution of sample proportions can be approximated by a normal distribution.



2. A *z* score associated with a sample proportion has a probability of $\alpha/2$ of falling in the right tail.



3. The *z* score separating the right-tail region is commonly denoted by $z_{\alpha/2}$ and is referred to as a ***critical value*** because it is on the borderline separating *z* scores from sample proportions that are likely to occur from those that are unlikely to occur.

Definition

A ***critical value*** is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely to occur. The number $z_{\alpha/2}$ is a critical value that is a *z* score with the property that it separates an area of $\alpha/2$ in the right tail of the standard normal distribution.

Notation for Critical Value

The critical value $z_{\alpha/2}$ is the positive *z* value that is at the vertical boundary separating an area of $\alpha/2$ in the right tail of the standard normal distribution. (The value of $-z_{\alpha/2}$ is at the vertical boundary for the area of $\alpha/2$ in the ***left tail***.) The subscript $\alpha/2$ is simply a reminder that the *z* score separates an area of $\alpha/2$ in the ***right tail*** of the standard normal distribution.

Definition

When data from a simple random sample are used to estimate a population proportion *p*, the ***margin of error***, denoted by ***E***, is the maximum likely difference (with probability 1 – α, such as 0.95) between the observed proportion $\hat{p}$ and the true value of the population proportion *p*. The margin of error *E* is also called the maximum error of the estimate and can be found by multiplying the critical value and the standard deviation of the sample proportions:

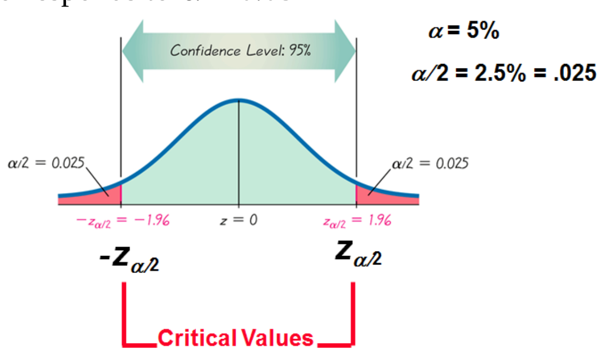
$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

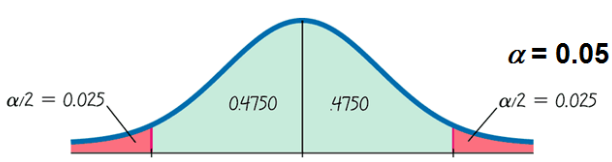
Example

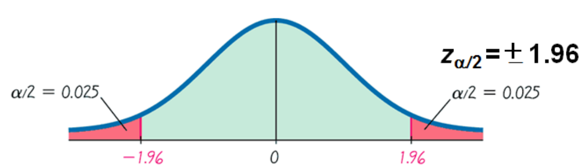
Find the critical value $z_{\alpha/2}$ corresponding to a 95% confidence level.

Solution

A 95% confidence level corresponds to $\alpha = 0.05$







The area in each of the red-shaded tails is $\frac{\alpha}{2} = 0.025$. The cumulative area to its left must be $1 - 0.024 = 0.975$. From the Normal Distribution Table, the area of 0.975 corresponds to $z = 1.96$. For a 95% confidence level, the critical value is therefore $z_{\alpha/2} = 1.96$

Confidence Interval for Estimating a Population Proportion *p*

Notation

p = population proportion

$\hat{p}$ = sample proportion

n = number of sample values

E = margin of error

$z_{\alpha/2}$ = z score separating an area of $\alpha/2$ in the right tail of the standard normal distribution

Requirements

1. The sample is a simple random sample.
2. The conditions for the binomial distribution are satisfied: there is a fixed number of trials, the trials are independent, there are two categories of outcomes, and the probabilities remain constant for each trial.
3. There are at least 5 successes and 5 failures.

Procedure for Constructing a Confidence Interval for *p*

1. Verify that the required assumptions are satisfied. (The sample is a simple random sample, the conditions for the binomial distribution are satisfied, and the normal distribution can be used to approximate the distribution of sample proportions because $np \geq 5$, and $nq \geq 5$ are both satisfied.)
2. Refer to Standard Normal Distribution Table and find the critical value $z_{\alpha/2}$ that corresponds to the desired confidence level.
3. Evaluate the margin of error $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$
4. Using the value of the calculated margin of error, E and the value of the sample proportion, $\hat{p}$, find the values of $\hat{p} - E$ and $\hat{p} + E$. Substitute those values in the general format for the confidence interval:

$$\hat{p} - E < p < \hat{p} + E$$

5. Round the resulting confidence interval limits to three significant digits.

Example

In the Chapter Problem we noted that a Pew Research Center poll of 1501 randomly selected U.S. adults showed that 70% of the respondents believe in global warming. The sample results are n = 1501, and $\hat{p} = 0.70$

a) Find the margin of error E that corresponds to a 95% confidence level.
b) Find the 95% confidence interval estimate of the population proportion p.
c) Based on the results, can we safely conclude that the majority of adults believe in global warming?
d) Assuming that you are a newspaper reporter, write a brief statement that accurately describes the results and includes all of the relevant information.

Solution

Requirement check: simple random sample; fixed number of trials, 1501; trials are independent; two categories of outcomes (believes or does not); probability remains constant. Note: number of successes and failures are both at least 5.

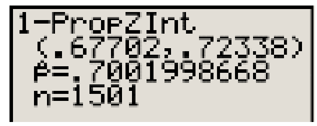
a) Use the formula to find the margin of error.

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= 1.96\sqrt{\frac{(0.70)(0.30)}{1501}}$$

$$= 0.023183$$

b) The 95% confidence interval:



$$\hat{p} - E < p < \hat{p} + E$$

$$0.70 - 0.023183 < p < 0.70 + 0.023183$$

$$0.677 < p < 0.723$$

***TI*-84:** Go STATS → TESTS → select A: 1-PropZInt

For *x* values mulptiply .7 (70%) by 1501 (*n*), however go back and round the number (.7*1501 = 1050.7) therefore the *x*-value is 1051

c) Based on the confidence interval obtained in part (*b*), it does appear that the proportion of adults who believe in global warming is greater than 0.5 (or 50%), so we can safely conclude that the majority of adults believe in global warming. Because the limits of **0.677** and **0.723** are likely to contain the true population proportion, it appears that the population proportion is a value greater than 0.5.

d) Here is one statement that summarizes the results: 70% of United States adults believe that the earth is getting warmer. That percentage is based on a Pew Research Center poll of 1501 randomly selected adults in the United States. In theory, in 95% of such polls, the percentage should differ by no more than 2.3 percentage points in either direction from the percentage that would be found by interviewing all adults in the United States.

Analyzing Polls

When analyzing polls consider:

1. The sample should be a simple random sample, not an inappropriate sample (such as a voluntary response sample).
2. The confidence level should be provided. (It is often 95%, but media reports often neglect to identify it.)
3. The sample size should be provided. (It is usually provided by the media, but not always.)
4. Except for relatively rare cases, the quality of the poll results depends on the sampling method and the size of the sample, but the size of the population is usually not a factor.

Caution

Never follow the common misconception that poll results are unreliable if the sample size is a small percentage of the population size. The population size is usually not a factor in determining the reliability of a poll.

Sample Size

Suppose we want to collect sample data in order to estimate some population proportion.
Sample size needed for a specified margin of error, E, and level of confidence $(1 - \alpha)$:

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$$

Round-Off Rule for Determining Sample Size

If the computed sample size n is not a whole number, round the value of n up to the next larger whole number.

Example

The Internet is affecting us all in many different ways, so there are many reasons for estimating the proportion of adults who use it. Assume that a manager for E-Bay wants to determine the current percentage of U.S. adults who now use the Internet. How many adults must be surveyed in order to be 95% confident that the sample percentage is in error by no more than three percentage points?

a) In 2006, 73% of adults used the Internet.

b) No known possible value of the proportion.

<u>Solution</u>

a) ***Given***: $\hat{p} = 0.73 \quad so \quad \hat{q} = 1 - 0.73 = 0.27$

With a 95% confidence level, we have $\alpha = 0.05$, so $z_{\alpha/2} = 1.96$. Also the margin of error is $E = 0.03$ (the decimal equivalent of "3 percentage points")

$$n = \frac{\left[z_{\alpha/2}\right]^2 \hat{p}\hat{q}}{E^2} = \frac{(1.96)^2(0.73)(0.27)}{0.03^2}$$

$\underline{\approx 842}$

We must obtain a simple random sample that includes at least 842 adults.

b) ***Given***: $z_{\alpha/2} = 1.96$ and $E = 0.03$

But with no prior knowledge of $\hat{p} \quad or \quad \hat{q}$

$$n = \frac{\left[z_{\alpha/2}\right]^2 \hat{p}\hat{q}}{E^2} = \frac{(1.96)^2(0.25)}{0.03^2}$$

$\underline{\approx 1068}$

To be 95% confident that our sample percentage is within three percentage points of the true percentage for all adults, we should obtain a simple random sample of 1068 adults.

Finding the Point Estimate and E from a Confidence Interval

Point estimate of p: $\hat{p} = \dfrac{(\text{upper confidence limit}) + (\text{lower confidence limit})}{2}$

Margin Error: $E = \dfrac{(\text{upper confidence limit}) - (\text{lower confidence limit})}{2}$

Suppose that a simple random sample of size n is taken from a population. A $(1 - \alpha)\cdot 100\%$ confidence interval for p is given by the following quantities

Lower bound: $\hat{p} - E \quad \to \quad \hat{p} - z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

Upper bound: $\hat{p} + E \quad \to \quad \hat{p} + z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

Note: It must be the case that $n\hat{p}\hat{q}$ and $n \le 0.05N$ to construct this interval.

Example

The article "High-Dose Nicotine Patch Therapy," by Dale Hurt, includes this statement: "off the 71 subjects, 70% were abstinent from smoking at 8 weeks (95% confidence interval, 58% to 81%)," Use that statement to find the point estimate $\hat{p}$ and the margin error E.

Solution

$$\hat{p} = \frac{(\text{upper confidence limit}) + (\text{lower confidence limit})}{2}$$
$$= \frac{0.81 + 0.58}{2}$$
$$= 0.695$$

$$E = \frac{(\text{upper confidence limit}) - (\text{lower confidence limit})}{2}$$
$$= \frac{0.81 - 0.58}{2}$$
$$= 0.115$$

Example

In July of 2008, a Quinnipiac University Poll asked 1783 registered voters nationwide whether they favored or opposed the death penalty for persons convicted of murder. 1123 were in favor.
Obtain a 90% confidence interval for the proportion of registered voters nationwide who are in favor of the death penalty for persons convicted of murder.

Solution

$$\hat{p} = 0.63$$

$$n\hat{p}\hat{q} = 1783(0.63)(1 - 0.63) = 415.6 \ge 0$$

The sample side is definitely less than 5% of the population size

$\alpha = 0.10 \quad \Rightarrow \quad z_{\alpha/2} = z_{0.05} = 1.645$

Lower bound: $\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.63 - 1.645\sqrt{\frac{0.63(1-0.63)}{1783}} \approx 0.61$

Upper bound: $\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.63 + 1.645\sqrt{\frac{0.63(1-0.63)}{1783}} \approx 0.65$

We are 90% confident that the proportion of registered voters who are in favor of the death penalty for those convicted of murder is between **0.61**and **0.65**.

Exercises *Section* 3.2 – Estimating a Population Proportion

1. Find the critical value $z_{\alpha/2}$ that corresponds to a 99% confidence level.

2. Find the critical value $z_{\alpha/2}$ that corresponds to a 99.5% confidence level.

3. Find the critical value $z_{\alpha/2}$ that corresponds to a 98% confidence level.

4. Find $z_{\alpha/2}$ for $\alpha = 0.10$.

5. Find $z_{\alpha/2}$ for $\alpha = 0.02$.

6. Express the confidence interval $0.200 < p < 0.500$ in the form $\hat{p} \pm E$

7. Express the confidence interval $0.42 < p < 0.54$ in the form $\hat{p} \pm E$

8. Express the confidence interval 0.222 ± 0.044 in the form $\hat{p} - E < p < \hat{p} + E$

9. Find the point estimate $\hat{p}$ and the margin of error E of $(0.320,\ 0.420)$

10. Find the margin of error E of $0.542 < p < 0.576$

11. Find the point estimate $\hat{p}$ of $0.824 < p < 0.868$

12. Find the point estimate $\hat{p}$ and the margin of error E of $0.772 < p < 0.776$

13. Find the point estimate $\hat{p}$ and the margin of error E of $0.433 < p < 0.527$

14. Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 1000, \quad x = 400, \quad 95\%\ \textit{confidence}$

15. Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 500, \quad x = 220, \quad 99\%\ \textit{confidence}$

16. Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given $n = 390, \quad x = 130, \quad 90\%\ \textit{confidence}$

17. Assume that a sample is used to estimate a population proportion p. Find the margin of error E that corresponds to the given 98% confidence; the sample size is 1230, of which 40% are successes.

18. Construct the confidence interval estimate of the population proportion p that corresponds to the given $n = 200, \quad x = 40, \quad 95\%\ \textit{confidence}$

19. Construct the confidence interval estimate of the population proportion p that corresponds to the given $n = 1236, \quad x = 109, \quad 99\%\ \textit{confidence}$

20. Construct the confidence interval estimate of the population proportion p that corresponds to the given $n = 5200, \quad x = 4821, \quad 99\%\ \textit{confidence}$

21. Find the minimum sample size requires to estimate a population proportion or percentage:
Margin of error: 0.045; confidence level: 95%: $\hat{p}\ \textit{and}\ \hat{q}\ \textit{unknown}$

22. Find the minimum sample size requires to estimate a population proportion or percentage: Margin of error: 2% points; confidence level: 99%: from prior study, $\hat{p}$ is estimate by the decimal equivalent of 14%

23. The Genetics and IVF Institute conducted a clinical trial of the XSORT method designed to increase the probability of conceiving a girl. As of this writing, 574 babies were born to parents using the XSORT method, and 525 of them were girls.

a) What is the best point estimate of the population proportion of girls born to parents using the XSORT method?

b) Use the sample data to construct a 95% confidence interval estimate of the percentage of girls born to parents using the XSORT method.

c) Based on the results, does the XSORT method appear to be effective? Why or why not?

24. An important issue facing Americans is the large number of medical malpractice lawsuits and the expenses that they generate. In a study of 1228 randomly selected medical malpractice lawsuits, it is found that 856 of them were later dropped or dismissed.

a) What is the best point estimate of the proportion of medical malpractice lawsuits that are dropped or dismissed?

b) Construct a 99% confidence interval estimate of the proportion of medical malpractice lawsuits that are dropped or dismissed.

c) Does it appear that the majority of such suits are dropped or dismissed?

25. A study of 420,095 Danish cell phone users found that 135 of them developed cancer was found to be 0.0340% for those not using cell phones.

a) Use the sample data to construct a 95% confidence interval estimate of the percentage of cell phone users who develop cancer of the brain or nervous system.

b) Do cell phone users appear to have a rate of cancer of the brain or nervous system that is different from the rate of such cancer among those not using cells phones? Why or why not?

26. In an Account survey of 150 senior executives, 47% said that the most common job interview mistake is to have little or no knowledge of the company. Construct a 99% confidence interval estimate of the proportion of all senior executives who have that same opinion. Is it possible that exactly half of all senior executives believe that the most common job interview mistake is to have little or no knowledge of the company? Why or why not?

Section 3.3 – Estimating a Population Mean

Point Estimate of the Population Mean

The sample mean $\overline{x}$ is the best ***point estimate*** of the population mean μ.

Confidence Interval for Estimating a Population Mean (with σ Known)

Objective

Construct a confidence interval used to estimate a population mean.

Notation

μ = population mean
σ = population standard deviation
$\overline{x}$ = sample mean
n = number of sample values
E = margin of error
$z_{\alpha/2} = z$ score separating an area of $\alpha/2$ in the right tail of the standard normal distribution

Requirements

1. The sample is a simple random sample. (All samples of the same size have an equal chance of being selected.)
2. The value of the population standard deviation σ is known.
3. Either or both of these conditions is/are satisfied: The population is normally distributed or $n > 30$.

Confidence Interval

$$\overline{x} - E < \mu < \overline{x} + E \quad \textit{wher} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\text{Or} \quad \overline{x} \pm E \qquad \text{Or} \quad \left(\overline{x} - E,\ \overline{x} + E\right)$$

Definition

The two values $\overline{x} - E$ and $\overline{x} + E$ are called ***confidence interval limits***.

Sample Mean

1. For all populations, the sample mean x is an unbiased estimator of the population mean μ, meaning that the distribution of sample means tends to center about the value of the population mean μ.
2. For many populations, the distribution of sample means x tends to be more consistent (with less variation) than the distributions of other sample statistics.

Procedure for Constructing a Confidence Interval for μ (with Known σ)

1. Verify that the requirements are satisfied.
2. Refer to Standard Normal Distribution Table or use technology to find the critical value $z_{\alpha/2}$ that corresponds to the desired confidence level.
3. Evaluate the margin of error $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
4. Find the values of $\bar{x} - E$ and $\bar{x} + E$. Substitute those values in the general format of the confidence interval: $\bar{x} - E < \mu < \bar{x} + E$
5. Round using the confidence intervals round-off rules.

Round-Off Rule for Confidence Intervals Used to Estimate μ

1. When using the *original set of data*, round the confidence interval limits to one more decimal place than used in original set of data.
2. When the original set of data is unknown and only the **summary statistics** (n, $\bar{x}$, s) are used, round the confidence interval limits to the same number of decimal places used for the sample mean.

Example

People have died in boat and aircraft accidents because an obsolete estimate of the mean weight of men was used. In recent decades, the mean weight of men has increased considerably, so we need to update our estimate of that mean so that boats, aircraft, elevators, and other such devices do not become dangerously overloaded. Using the weights of men from Data Set 1 in Appendix B, we obtain these sample statistics for the simple random sample: $n = 40$ and $\bar{x} = 172.55$ lb. Research from several other sources suggests that the population of weights of men has a standard deviation given by $\sigma = 26$ lb.

a) Find the best point estimate of the mean weight of the population of all men.

b) Construct a 95% confidence interval estimate of the mean weight of all men.

c) What do the results suggest about the mean weight of 166.3 lb that was used to determine the safe passenger capacity of water vessels in 1960 (as given in the National Transportation and Safety Board safety recommendation M-04-04)?

Solution

a) The sample mean of 172.55 lb. is the best point estimate of the mean weight of the population of all men

b) A 95% confidence interval or 0.95 implies $\sigma = 0.05$, so $z_{\alpha/2} = 1.96$.

Calculate the margin of error.

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{26}{\sqrt{40}}$$

$$= 8.0574835$$

The confidence interval: $\bar{x} - E < \mu < \bar{x} + E$

$$172.55 - 8.0574835 < \mu < 172.55 + 8.0574835$$

$$164.49 < \mu < 180.61$$

c) Based on the confidence interval, it is possible that the mean weight of 166.3 lb. used in 1960 could be the mean weight of men today. However, the best point estimate of 172.55 lb. suggests that the mean weight of men is now considerably greater than 166.3 lb. considering that an underestimate of the mean weight of men could result in lives lost through overloaded boats and aircraft, these results strongly suggest that additional data should be collected. (Additional data have been collected, and the assumed mean weight of men has been increased.)

Finding a Sample Size for Estimating a Population Mean

μ = population mean

σ = population standard deviation

$\overline{x}$ = population standard deviation

E = desired margin of error

$z_{\alpha/2}$ = z score separating an area of $\alpha/2$ in the right tail of the standard normal distribution

$$n = \left[\frac{z_{\alpha/2} \bullet \sigma}{E}\right]^2$$

Finding the Sample Size *n* when σ is Unknown

1. Use the range rule of thumb to estimate the standard deviation as follows: $\sigma \approx$ range/4.
2. Start the sample collection process without knowing σ and, using the first several values, calculate the sample standard deviation s and use it in place of σ. The estimated value of σ can then be improved as more sample data are obtained, and the sample size can be refined accordingly.
3. Estimate the value of σ by using the results of some other study that was done earlier.

Example

Assume that we want to estimate the mean IQ score for the population of statistics students. How many statistics students must be randomly selected for IQ tests if we want 95% confidence that the sample mean is within 3 IQ points of the population mean?

Solution

For 95% confidence interval, we have $\alpha = 0.05 \quad so \quad z_{\alpha/2} = 1.96$

Since we need the sample mean to be within 3 IQ points of μ, the margin of error is E = 3. Also, $\sigma = 15$

$$n = \left[\frac{z_{\alpha/2} \bullet \sigma}{E}\right]^2 = \left[\frac{1.96 \bullet 15}{3}\right]^2$$

≈ 97

With a simple random sample of only 97 statistics students, we will be 95% confident that the sample mean is within 3 IQ points of the true population mean μ.

Student *t* Distribution

Definitions

Suppose that a simple random sample of size *n* is taken from a population. If the population from which the sample is drawn follows a normal distribution, the distribution of

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

follows Student's *t*-distribution with $n - 1$ degrees of freedom where $\bar{x}$ is the sample mean and *s* is the sample standard deviation.

Distribution for all samples of size *n*. It is often referred to as a *t* distribution and is used to find criti***cal values*** denoted by $t_{\alpha/2}$.

The number of ***degrees of freedom*** (*df*).for a collection of sample data is the number of sample values that can vary after certain restrictions have been imposed on all data values.

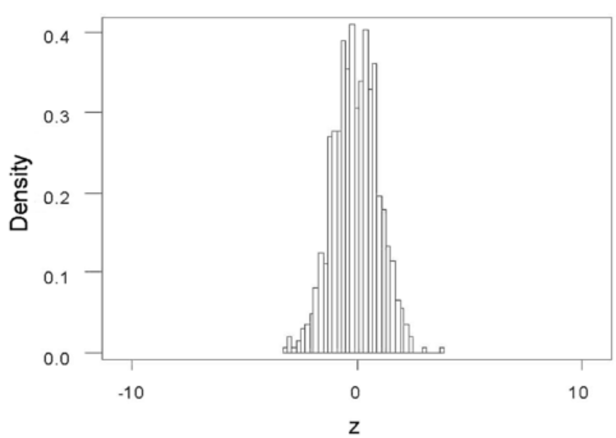
Degrees of freedom: *df* = *n* – 1

Example

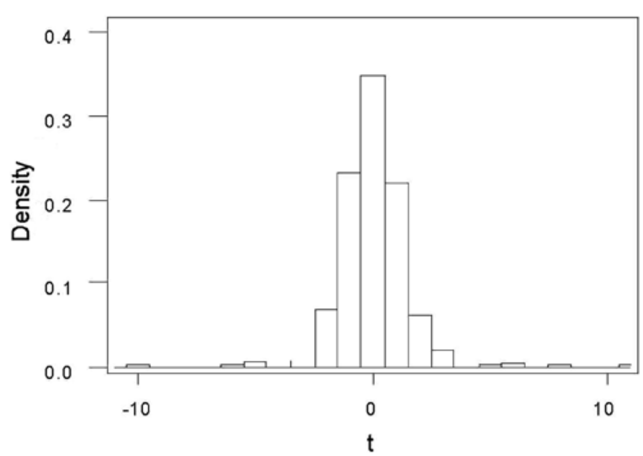
Obtain 1,000 simple random samples of size $n = 5$ from a normal population with $\mu = 50$ and $\sigma = 10$. Determine the sample mean and sample standard deviation for each of the samples.

Compute $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$ and $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ for each sample. Draw a histogram for both *z* and *t.*

Solution



Histogram for z

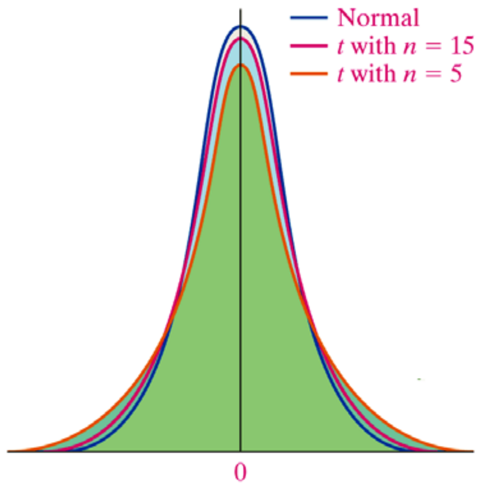


Histogram for t

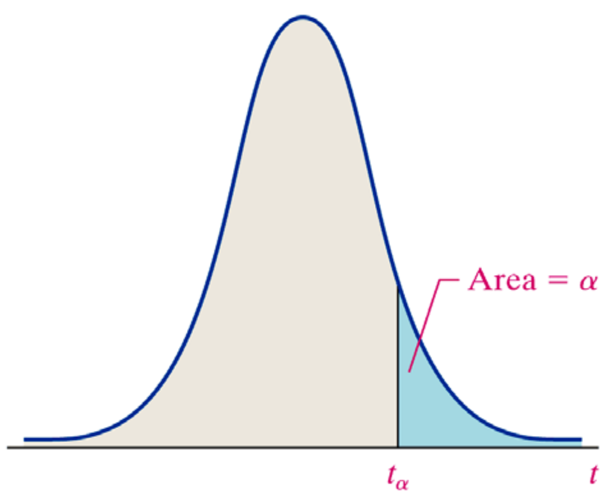
- ✓ The histogram for *z* is symmetric and bell-shaped with the center of the distribution at 0 and virtually all the rectangles between –3 and 3. In other words, *z* follows a standard normal distribution
- ✓ The histogram for *t* is also symmetric and bell-shaped with the center of the distribution at 0, but the distribution of *t* has longer tails (i.e., *t* is more dispersed), so it is unlikely that *t* follows a standard normal distribution. The additional spread in the distribution of *t* can be attributed to the fact that we use *s* to find *t* instead of σ. Because the sample standard deviation is itself a random variable (rather than a constant such as σ), we have more dispersion in the distribution of *t*.

Properties of the *t*-Distribution

1. The *t*-distribution is different for different degrees of freedom.
2. The *t*-distribution is centered at 0 and is symmetric about 0.
3. The area under the curve is 1. The area under the curve to the right of 0 equals the area under the curve to the left of 0, which equals 1/2.
4. As t increases or decreases without bound, the graph approaches, but never equals, zero.
5. The area in the tails of the *t*-distribution is a little greater than the area in the tails of the standard normal distribution, because we are using s as an estimate of σ, thereby introducing further variability into the *t*- statistic.
6. As the sample size n increases, the density curve of t gets closer to the standard normal density curve. This result occurs because, as the sample size n increases, the values of s get closer to the values of σ, by the Law of Large Numbers.



The notation t_α is the ***t*-value** such that the area under the standard normal curve to the right is α.



Example

Find the t-value such that the area under the t-distribution to the right of the t-value is 0.2 assuming 10 degrees of freedom. That is, find $t_{0.20}$ with 10 degrees of freedom.

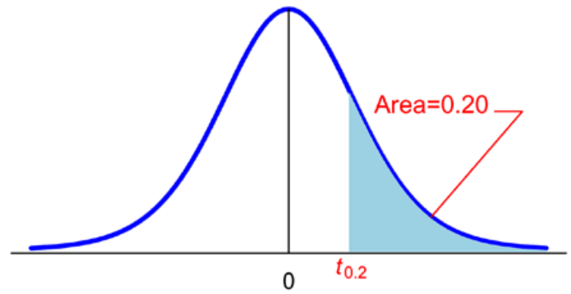
Solution

t distribution critical values

	Upper-tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587

$df = 10$

$t_{0.20} = 0.879$



Example

A sample of size $n = 7$ is a simple random sample selected from a normally distributed population. Find the critical value $t_{\alpha/2}$ corresponding to a 95% confidence level.

Solution

Because $n = 7$, the number of degrees of freedom is: $n - 1 = 6$.

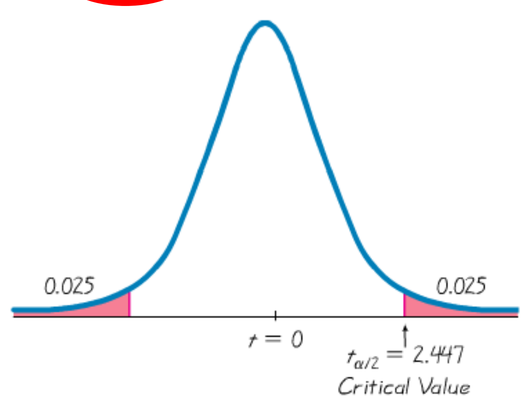
Using t–Distribution Table:

Degrees of Freedom	Area in Two Tails				
	0.01	0.02	0.05	0.10	0.20
6	**3.707**	3.143	**2.447**	1.943	1.440

The value is 2.447.

$t_{\alpha/2} = t_{0.05/2} = t_{0.025} = 2.447$

Such critical values $t_{\alpha/2}$ are used for the margin of error E and confidence interval.



Margin of Error E for Estimate of m (With σ Not Known)

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$$

where $t_{\alpha/2}$ has n – 1 degrees of freedom.

Notation

μ = population mean
$\overline{x}$ = sample mean
s = sample standard deviation
n = number of sample values
E = margin of error
$t_{\alpha/2}$ = critical t value separating an area of $\alpha/2$ in the right tail of the t distribution

Confidence Interval for the Estimate of μ (With σ Not Known)

$$\overline{x} - E < \mu < \overline{x} + E$$

Where $E = t_{\alpha/2} \frac{s}{\sqrt{n}}$ $\quad df = n - 1$

Procedure for Constructing a Confidence Interval for μ (With σ Unknown**)**

1. Verify that the requirements are satisfied.
2. Using n – 1 degrees of freedom, refer to Table A-3 or use technology to find the critical value $t_{\alpha/2}$ that corresponds to the desired confidence level.
3. Evaluate the margin of error $E = t_{\alpha/2} \frac{s}{\sqrt{n}}$
4. Find the values of $\overline{x} - E < \mu < \overline{x} + E$. Substitute those values in the general format for the confidence interval:
5. Round the resulting confidence interval limits.

Example

A common claim is that garlic lowers cholesterol levels. In a test of the effectiveness of garlic, 49 subjects were treated with doses of raw garlic, and their cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have a mean of 0.4 and a standard deviation of 21.0. Use the sample statistics of n = 49, $\overline{x}$ = 0.4 and s = 21.0 to construct a 95% confidence interval estimate of the mean net change in LDL cholesterol after the garlic treatment. What does the confidence interval suggest about the effectiveness of garlic in reducing LDL cholesterol?

Solution

Requirements are satisfied: simple random sample and n = 49 (i.e., n > 30).
95% implies a = 0.05.
With n = 49, the df = 49 – 1 = 48

Closest df is 50, two tails, so $t_{\alpha/2} = 2.009$

Using $t_{\alpha/2} = 2.009$, $s = 21.0$ and $n = 49$ the margin of error is:

$$E = t_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$= 2.009 \cdot \frac{21.0}{\sqrt{49}}$$

$$= 6.027$$

Construct the confidence interval: $\bar{x} = 0.4$; $E = 6.027$

$$\bar{x} - E < \mu < \bar{x} + E$$

$$0.4 - 6.027 < \mu < 0.4 + 6.027$$

$$-5.6 < \mu < 6.4$$

We are 95% confident that the limits of –5.6 and 6.4 actually do contain the value of μ, the mean of the changes in LDL cholesterol for the population. Because the confidence interval limits contain the value of 0, it is very possible that the mean of the changes in LDL cholesterol is equal to 0, suggesting that the garlic treatment did not affect the LDL cholesterol levels. It does not appear that the garlic treatment is effective in lowering LDL cholesterol.

Example

Construct a 99% confidence interval about the population mean weight (in grams) of pennies minted after 1982. Assume μ = 0.02 grams.

2.46	2.47	2.49	2.48	2.5	2.44	2.46	2.45	2.49
2.47	2.45	2.46	2.45	2.46	2.47	2.44	2.45	

Solution

$$t_{\alpha/2} = t_{.01/2} = t_{.005} = 1.746$$

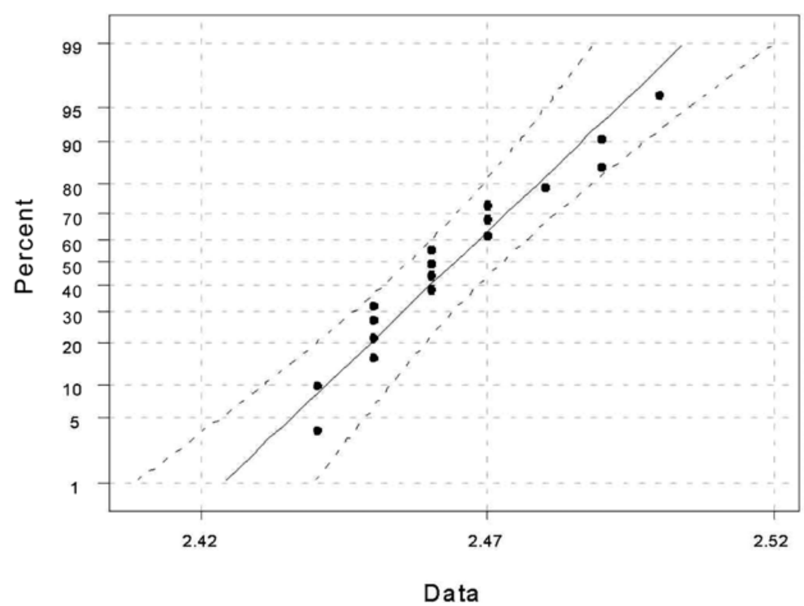
$$E = t_{\alpha/2}\frac{s}{\sqrt{n}} = 1.746\frac{.02}{\sqrt{17}} = .008$$

$$\bar{x} - E < \mu < \bar{x} + E$$

$$2.464 - .008 < \mu < 2.464 + .008$$

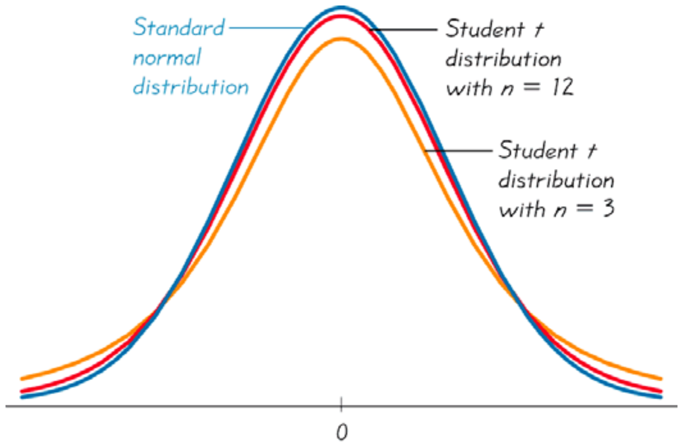
$$2.456 < \mu < 2.472$$

We are 99% confident that the mean weight of pennies minted after 1982 is between 2.456 and 2.472 grams.



Important Properties of the Student *t* Distribution

1. The Student *t* distribution is different for different sample sizes (see the following slide, for the cases $n = 3$ and $n = 12$).



2. The Student *t* distribution has the same general symmetric bell shape as the standard normal distribution but it reflects the greater variability (with wider distributions) that is expected with small samples.
3. The Student *t* distribution has a mean of $t = 0$ (just as the standard normal distribution has a mean of $z = 0$).
4. The standard deviation of the Student *t* distribution varies with the sample size and is greater than 1 (unlike the standard normal distribution, which has a $s = 1$).
5. As the sample size *n* gets larger, the Student *t* distribution gets closer to the normal distribution.

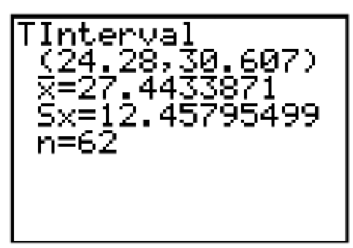
Finding the Point Estimate and *E* from a Confidence Interval

Point estimate of μ: $$\bar{x} = \frac{(upper\ confidence\ limit) + (lower\ confidence\ limit)}{2}$$

Margin of Error: $$E = \frac{(upper\ confidence\ limit) - (lower\ confidence\ limit)}{2}$$

TI-83/84 PLUS The TI-83/84 Plus calculator can be used to generate confidence intervals for original sample values stored in a list, or you can use the summary statistics n, $\bar{x}$, and s. Either enter the data in list L1 or have the summary statistics available, then press the **STAT** key. Now select **TESTS** and choose **TInterval** if σ is not known. (Choose **ZInterval** if σ is known.) After making the required entries, the calculator display will include the confidence interval in the format of $(\bar{x} - E, \bar{x} + E)$.

TI-83/84 PLUS



Confidence Intervals for Comparing Data

Confidence intervals can be used informally to compare different data sets, but the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of means.

Determining the Sample Size *n*

The sample size required to estimate the population mean, μ, with a level of confidence $(1-\alpha)\cdot 100\%$ with a specified margin of error, E, is given by

$$n = \left[\frac{z_{\alpha/2} \bullet s}{E}\right]^2$$

where n is rounded up to the nearest whole number.

Example

How large a sample would be required to estimate the mean weight of a penny manufactured after 1982 within 0.005 grams with 99% confidence? Assume $\sigma = 0.02$.

Solution

Given: $s = 0.02 \quad E = 0.005$

$z_{\alpha/2} = z_{0.005} = 2.575$

$$n = \left(\frac{z_{\alpha/2} \bullet s}{E}\right)^2 = \left(\frac{(2.575)(0.02)}{0.005}\right)^2 \underline{=106.09}$$

$\therefore \underline{n = 107}$

Exercises *Section* 3.3 – Estimating a Population Mean

1. A design engineer of the Ford Motor Company must estimate the mean leg length of all adults. She obtains a list of the 1275 employees at her facility; then obtains a simple random sample of 50 employees. If she uses this sample to construct a 95% confidence interval to estimate the mean leg length for the population of all adults, will her estimate be good? Why or why not?

2. Find the critical value $z_{\alpha/2}$ that corresponds to a 90% confidence level.

3. Find the critical value $z_{\alpha/2}$ that corresponds to a 98% confidence level.

4. Find $z_{\alpha/2}$ for $\alpha = 0.20$

5. Find $z_{\alpha/2}$ for $\alpha = 0.07$

6. How many adults must be randomly selected to estimate the mean FICO (credit rating) score of working adults in U.S.? We want 95% confidence that the sample mean is within 3 points of the population mean, and the population standard deviation is 68.

7. A simple random sample of 40 salaries of NCAA football coaches has a mean of \$415,953. Assume that σ = \$463,364.

a) Find the best estimate of the mean salary of all NCAA football coaches.
b) Construct a 95% confidence interval estimate of the mean salary of an NCAA football coach.
c) Does the confidence interval contain the actual population mean of \$474,477?

8. A simple random sample of 50 adults (including males and females) is obtained, and each person's red blood cell count (in cells per microliter) is measured. The sample mean is 4.63. The population standard deviation for red blood cell counts is 0.54.

a) Find the best point estimate of the mean red blood cell count of adults.
b) Construct a 99% confidence interval estimate of the mean red blood cell count of adults.
c) The normal range of red blood cell counts for adults is 4.7 to 6.1 for males and 4.3 to 5.4 for females. What does the confidence interval suggest about these normal ranges?

9. A simple random sample of 125 SAT scores has a mean of 1522. Assume that SAT scores have a standard deviation of 333.

a) Construct a 95% confidence interval estimate of the mean SAT score.
b) Construct a 99% confidence interval estimate of the mean SAT score.
c) Which of the preceding confidence intervals is wider? Why?

10. When 14 different second-year medical students measured the blood pressure of the same person, they obtained the results listed below. Assuming that the population standard deviation is known to be 10 mmHg, construct a 95% confidence interval estimate of the population mean. Ideally, what should the confidence interval be in this situation?

138 130 135 140 120 125 120 130 130 144 143 140 130 150

11. Do the given conditions justify using the margin of error $E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$ when finding a confidence interval estimate of the population mean μ?

a) The sample size is $n = 4$, $\sigma = 12.5$, and the original population is normally distributed

b) The sample size is $n = 5$ and σ is not known

12. Use the confidence level and sample data to find the margin of error E.

a) Replacement times for washing machines: 90% confidence; $n = 37$, $\bar{x} = 10.4\ yrs$, $\sigma = 2.2\ yrs$

b) College students' annual earnings: 99% confidence; $n = 76$, $\bar{x} = \$4196$, $\sigma = \$848$

13. Use the confidence level and sample data to find a confidence interval for estimating the population μ. A laboratory tested 89 chicken eggs and found that the mean amount of cholesterol was 203 milligrams with $\sigma = 11.4\ mg$. Construct a 95% confidence interval for the true mean cholesterol content μ, of all such eggs.

14. Use the confidence level and sample data to find a confidence interval for estimating the population μ. A group of 66 randomly selected students have a mean score of 34.3 on a placement test. The population standard deviation $\sigma = 3$. What is the 90% confidence interval for the mean score, μ, of all students taking the test?

15. Use the given information to find the minimum sample size required to estimate an unknown population mean μ. Margin error: \$139, confidence level: 99%, σ = \$522

16. What does it mean when we say that the methods for constructing confidence intervals in this section are robust against departures from normality? Are the methods for constructing confidence intervals in this section robust against poor sampling methods?

17. Assume that we want tic instruct a confidence interval using the given confidence level 95%; $n = 23$; $\boldsymbol{\sigma}$ is unknown; population appears to be normally distributed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

18. Assume that we want tic instruct a confidence interval using the given confidence level 99%; $n = 25$; $\boldsymbol{\sigma}$ is known; population appears to be normally distributed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

19. Assume that we want tic instruct a confidence interval using the given confidence level 99%; $n = 6$; σ is unknown; population appears to be very skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

20. Assume that we want tic instruct a confidence interval using the given confidence level 90%; $n = 200$; σ = 15.0; population appears to be skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

21. Assume that we want tic instruct a confidence interval using the given confidence level 95%; $n = 9$; σ is unknown; population appears to be very skewed. Do one of the following

a) Find the critical value $z_{\alpha/2}$

b) Find the critical value $t_{\alpha/2}$

c) State that neither the normal nor the *t*-distribution applies.

22. Given 95% *confidence*; $n = 20, \quad \bar{x} = \$9004, \quad s = \$569$. Assume that the sample is a simple random and the population has a normal distribution.

a) Find the margin error

b) Find the confidence interval for the population mean μ.

23. Given 99% *confidence*; $n = 7, \quad \bar{x} = 0.12, \quad s = 0.04$. Assume that the sample is a simple random and the population has a normal distribution.

a) Find the margin error

b) Find the confidence interval for the population mean μ.

24. In a test of the effectiveness of garlic for lowering cholesterol, 47 subjects were treated with Garlicin, which is garlic in a processed tablet form. Cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have a mean of 3.2 and standard deviation of 18.6.

a) What is the best point estimate of the population mean net change LDL cholesterol after Garlicin treatment?

b) Construct a 95% confidence interval estimate of the mean net change in LDL cholesterol after the Garlicin treatment. What does the confidence interval suggest about the effectiveness of Garlicin in reducing LDL cholesterol?

25. A random sample of the birth weights of 186 babies has a mean of 3103 g and a standard deviation of 696 g. These babies were born to mothers who did not use cocaine during their pregnancies.

a) What is the best point estimate of the mean weight of babies born to mothers who did not use cocaine during their pregnancies?

b) Construct a 95% confidence interval estimate of the mean birth for all such babies.

c) Compare the confidence interval from part (*b*) to this confidence interval obtained from birth weights of babies born to mothers who used cocaine during pregnancy: $2608\ g < \mu < 2792\ g$. Does cocaine use appear to affect the birth weight of a baby?

26. In a study designed to test the effectiveness of acupuncture for treating migraine, 142 subjects were treated with acupuncture and 80 subjects were given a sham treatment. The numbers of migraine attacks for the acupuncture treatment group had a mean of 1.8 and a standard deviation of 1.4. The numbers of migraine attacks for the sham treatment group had a mean of 1.6 and standard deviation of 1.2

a) Construct a 95% confidence interval estimate of the mean number of migraine attacks for those treated with acupuncture.

b) Construct a 95% confidence interval estimate of the mean number of migraine attacks for those given a sham treatment.

c) Compare the two confidence intervals. What do the results about the effectiveness of acupuncture?

27. 30 randomly selected students took the statistics final. If the sample mean was 79 and the standard deviation was 14.5, construct a 99% confidence interval for the mean score of all students.
Use the given degree of confidence and sample data to construct a confidence level interval for the population mean μ. Assume that the population has a normal distribution.

Section 3.4 – Estimating a Population Standard Deviation

Chi-Square Distribution

In a normally distributed population with variance σ^2 assume that we randomly select independent samples of size *n* and, for each sample, compute the sample variance s^2 (which is the square of the sample standard deviations). The sample statistic χ^2 (pronounced chi-square) has a sampling distribution called the ***chi-square distribution***.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Where

n = sample size

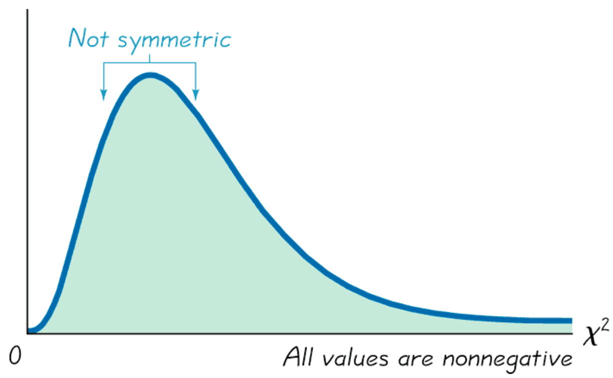
s^2 = sample variance

σ^2 = population variance

degrees of freedom = $n - 1$

Properties of the Distribution of the Chi-Square Statistic

1. The chi-square distribution is not symmetric, unlike the normal and Student *t* distributions. As the number of degrees of freedom increases, the distribution becomes more symmetric.



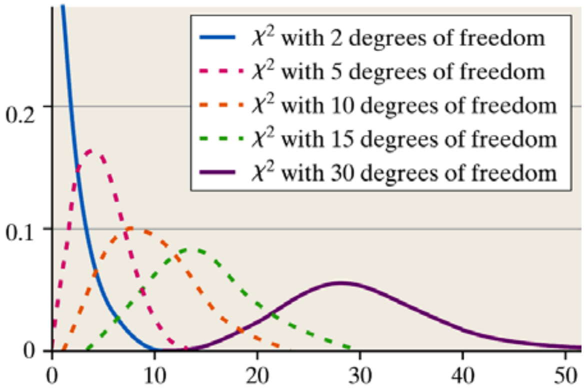
Chi-Square Distribution

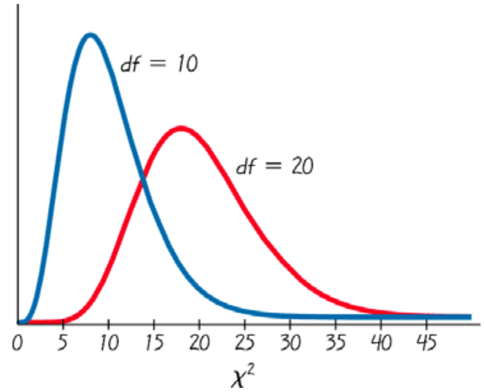
2. /The shape of the chi-square distribution depends on the degrees of freedom, just like the Student's *t*-distribution.

3. The values of chi-square can be zero or positive, but they cannot be negative.

4. The chi-square distribution is different for each number of degrees of freedom, which is df = $n - 1$. As the number of degrees of freedom increases, the chi-square distribution becomes more nearly symmetric.

In ***Table Chi-Square*** $\left(\chi^2\right)$ ***Distribution***, each critical value of χ^2 corresponds to an area given in the top row of the table, and that area represents the cumulative area located to the right of the critical value.

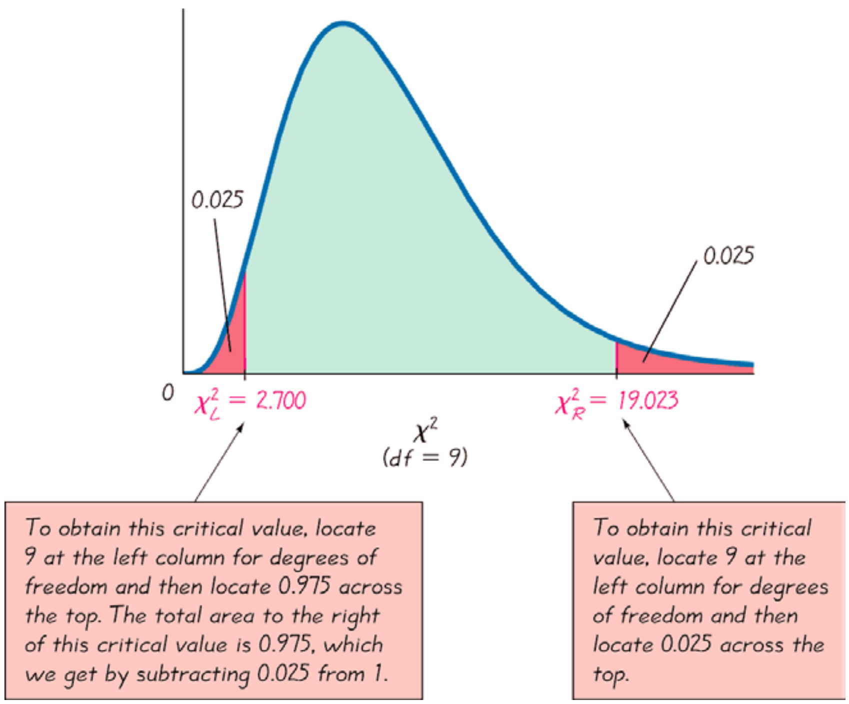




Chi-Square Distribution for df = 10 and df = 20

Example

A simple random sample of ten voltage levels is obtained. Construction of a confidence interval for the population standard deviation σ requires the left and right critical values of χ^2 corresponding to a confidence level of 95% and a sample size of n = 10. Find the critical value of χ^2 separating an area of 0.025 in the left tail, and find the critical value of χ^2 separating an area of 0.025 in the right tail.



Example

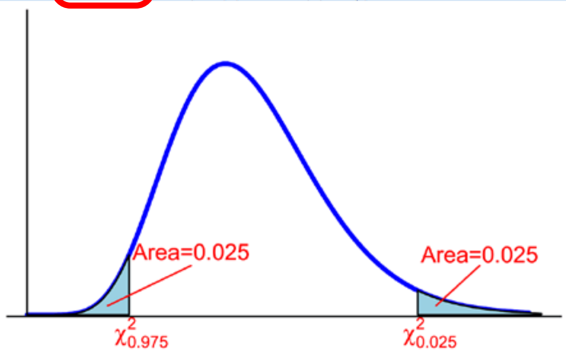
Find the chi-square values that separate the middle 95% of the distribution from the 2.5% in each tail. Assume 18 degrees of freedom.

Solution

Chi-Square (χ^2) Distribution

Area to the Right of Critical Value

Degrees of Freedom	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156



$\chi^2_{0.975} = 8.231$ $\qquad$ $\chi^2_{0.025} = 31.526$

Confidence Interval for Estimating a Population Standard Deviation or Variance

σ = population standard deviation
s = sample standard deviation
n = number of sample values
$\chi^2_L = \chi^2_{\alpha/2}$ = left-tailed critical value of χ^2

σ^2 = population variance
s^2 = sample variance
E = margin of error
$\chi^2_R = \chi^2_{1-\alpha/2}$ = right-tailed critical value of χ^2

Estimators of σ^2

The sample variance s^2 is the best point estimate of the population variance σ^2.

Estimators of σ

The sample standard deviation s is a commonly used point estimate of σ (even though it is a biased estimate).

Requirements:

1. The sample is a simple random sample.
2. The population must have normally distributed values (even if the sample is large).

Confidence Interval for the Population Variance σ^2

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

Confidence Interval for the Population Standard Deviation σ

$$\sqrt{\frac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_L}}$$

Procedure for Constructing a Confidence Interval for σ or σ^2

1. Verify that the required assumptions are satisfied.
2. Using $n - 1$ degrees of freedom, refer to ***Table Chi-Square*** $\left(\chi^2\right)$ ***Distribution*** or use technology to find the critical values χ^2_R and χ^2_L that correspond to the desired confidence level.
3. Evaluate the upper and lower confidence interval limits using this format of the confidence interval:
$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$
4. If a confidence interval estimate of σ is desired, take the square root of the upper and lower confidence interval limits and change σ^2 to σ.
5. Round the resulting confidence level limits. If using the original set of data to construct a confidence interval, round the confidence interval limits to one more decimal place than is used for the original set of data. If using the sample standard deviation or variance, round the confidence interval limits to the same number of decimals places.

Confidence Intervals for Comparing Data *Caution*

Confidence intervals can be used ***informally*** to compare the variation in different data sets, but *the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of variances or standard deviations*.

Example

The proper operation of typical home appliances requires voltage levels that do not vary much. Listed below are ten voltage levels (in volts) recorded in the author's home on ten different days. These ten values have a standard deviation of $s = 0.15$ volt. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all voltage levels.

123.3 123.5 123.7 123.4 123.6 123.5 123.5 123.4 123.6 123.8

Solution

Requirements are satisfied: simple random sample and normality

Construct the confidence interval: $n = 10$, $s = 0.15$

$$\chi_L^2 = 2.700 \quad and \quad \chi_R^2 = 19.023$$

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

$$\frac{(10-1)(0.15)^2}{19.023} < \sigma^2 < \frac{(10-1)(0.15)^2}{2.70}$$

$$0.010645 < \sigma^2 < 0.0750$$

Finding the square root of each part (before rounding), then rounding to two decimal places, yields this 95% confidence interval estimate of the population standard deviation:

$$0.10\ volt < \sigma < 0.27\ volt$$

✓ Based on this result, we have 95% confidence that the limits of 0.10 volt and 0.27 volt contain the true value of σ.

Example

We want to estimate the standard deviation σ of all voltage levels in a home. We want to be 95% confident that our estimate is within 20% of the true value of σ. How large should the sample be? Assume that the population is normally distributed.

Solution

From the Table, we can see that 95% confidence and an error of 20% for σ correspond to a sample of size 48. We should obtain a simple random sample of 48 voltage levels form the population of voltage levels.

TI-83/84 PLUS The TI-83/84 Plus calculator does not provide confidence intervals for σ or σ^2 directly, but the program **S2INT** can be used. That program was written by Michael Lloyd of Henderson State University, and it can be downloaded from www.aw.com/triola. The program S2INT uses the program ZZINEWT, so that program must also be installed. After storing the programs on the calculator, press the **PRGM** key, select **S2INT,** and enter the sample variance s^2, the sample size n, and the confidence level (such as 0.95). Press the **ENTER** key, and wait a while for the display of the confidence interval limits for σ^2. Find the square root of the confidence interval limits if an estimate of σ is desired.

Exercises Section 3.4 – Estimating a Population Standard Deviation

1. Using the weights of the M&M candies. We use the standard deviation of the sample $(s = 0.05179\ g)$ to obtain the following 95% confidence interval estimate of the standard deviation of the weights of all M&Ms: $0.0455\ g < \sigma < 0.0602\ g$. Write a statement that correctly interprets that confidence interval.

2. Find χ_L^2 and χ_R^2 that corresponds to: 95%; $n = 9$

3. Find χ_L^2 and χ_R^2 that corresponds to: 99%; $n = 81$

4. Find χ_L^2 and χ_R^2 that corresponds to: 90%; $n = 51$

5. Find a confidence interval for the population standard deviation σ.
95% *confidence*; $n = 30$, $\overline{x} = 1533$, $s = 333$ (*Assume has a normal distribution*)

6. Find a confidence interval for the population standard deviation σ
95% *confidence*; $n = 25$, $\overline{x} = 81.0\ mi/h$, $s = 2.3\ mi/h$ (*Assume has a normal distribution*)

7. Find a confidence interval for the population standard deviation σ
99% *confidence*; $n = 7$, $\overline{x} = 7.106$, $s = 2.019$ (*Assume has a normal distribution*)

8. In a study of the effects of prenatal cocaine use on infants, the following sample data were obtained for weights at birth: $n = 190$, $\overline{x} = 2700\ g$, $s = 645\ g$. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all birth weights of infants born to mothers who used cocaine during pregnancy. Because from the Table, a maximum of 100 degrees of freedom while we require 189 degrees of freedom, use these critical values to obtained $\chi_L^2 = 152.8222$ and $\chi_R^2 = 228.9638$. Based on the result, does the standard deviation appear to be different from the standard deviation of 696*g* for birth weights of babies born to mothers who did not use cocaine during pregnancy?

9. In the course of designing theather seats, the sitting heights (in *mm*) of a simple random sample of adults women is obtained, and the results are

849 807 821 856 864 877 772 848 802 807 887 815

Use the sample data to construct a 95% confidence interval estimate of $\boldsymbol{\sigma}$, the standard deviation of sitting heights of all women. Does the confidence contain the value of 35 *mm*, which is believed to be the standard deviation of sitting heights of women?

10. One way to measure the risk of a stock is through the standard deviation rate of return of the stock. The following data represent the weekly rate of return (in percent) of Microsoft for 15 randomly selected weeks. Compute the 90% confidence interval for the risk of Microsoft stock.

5.34	9.63	–2.38	3.54	–8.76	2.12	–1.95	0.27
0.15	5.84	–3.90	–3.80	2.85	–1.61	–3.31	

Section 3.5 – Language of Hypothesis Testing

If, under a given assumption, the probability of a particular observed event is exceptionally small, we conclude that the assumption is probably not correct.

A ***hypothesis*** is a statement regarding a characteristic of one or more populations.

Example

In 2008, 62% of American adults regularly volunteered their time for charity work. A researcher believes that this percentage is different today

Solution

According to a study published in March, 2006 the mean length of a phone call on a cellular telephone was 3.25 minutes. A researcher believes that the mean length of a call has increased since then.

Using an old manufacturing process, the standard deviation of the amount of wine put in a bottle was 0.23 ounces. With new equipment, the quality control manager believes the standard deviation has decreased.

Caution

We test these types of statements using sample data because it is usually impossible or impractical to gain access to the entire population. If population data are available, there is no need for inferential statistics.

Example

ProCare Industries, Ltd provided a product called "Gender Choice", which, according to advertising claims, allowed couples to "increase your chances of having a girl up to 80%." Suppose we conduct an experiment with 100 couples who want to have baby girls, and they all follow the Gender Choice "easy-to-use in-home system" described in the pink package designed for girls. Assuming that Gender Choice has no effect and using only common sense and no formal statistical methods, what should we conclude about the assumption of "no effect" from Gender Choice if 100 couples using Gender Choice have 100 babies consisting of the following

a) 52 girls

b) 97 girls

Solution

a) We normally expect around 50 girls in 100 births. The result of 52 girls is close to 50, so the Gender Choice product is effective. The result of 52 girls could easily occur by chance, so there isn't sufficient evidence to say that Gender Choice is effective, even though the sample proportion of girls is greater than 50%.

b) The result of 97 girls in 100 births is extremely unlikely to occur by chance.
Either an extremely rare event has occurred by chance, or Gender Choice is effective. The extremely low probability of getting 97 girls suggests that Gender Choice is effective.

Using Statistics

- A hypothesis is a statement or assertion about the state of nature (about the true value of an unknown population parameter):
  - ✓ The accused is innocent
  - ✓ $\mu = 100$
- Every hypothesis implies its contradiction or alternative
  - ✓ The accused is guilty
  - ✓ $\mu \neq 100$
- A hypothesis is either true or false, and you may fail to reject it or you may reject it on the basis of information
  - ✓ Trial testimony and evidence
  - ✓ Sample data

Making Decision

One hypothesis is maintained to be true until a decision is made to reject it as false:

- ✓ Guilt is proven "beyond a reasonable doubt"
- ✓ The alternative is highly improbable

A decision to fail to reject or reject a hypothesis may be:

- ✓ Correct
  A true hypothesis may not be rejected
  » An innocent defendant may be acquitted
  A false hypothesis may be rejected
  » A guilty defendant may be convicted
- ✓ Incorrect
  A true hypothesis may be rejected
  » An innocent defendant may be convicted
  A false hypothesis may not be rejected
  » A guilty defendant may be acquitted

Components of a Formal Hypothesis Test

Hypothesis testing is a procedure, based on sample evidence and probability, used to test statements regarding a characteristic of one or more populations.

Null Hypothesis: The ***null hypothesis***, denoted H_0, is a statement to be tested. The null hypothesis is a statement of no change, no effect or no difference and is assumed true until evidence indicates otherwise.

- The null hypothesis (denoted by H_0) is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value.

- We test the null hypothesis directly.
- Either reject H_0 or fail to reject H_0.
- $H_0 : \mu = 100$

Alternative Hypothesis: The ***alternative hypothesis***, denoted H_1, is a statement that we are trying to find evidence to support.

- The alternative hypothesis (denoted by H_1 or H_a or H_A) is the statement that the parameter has a value that somehow differs from the null hypothesis.
- The symbolic form of the alternative hypothesis must use one of these symbols: ≠, <, >.
- $H_1 : \mu \neq 100$

Three ways to set up the null and alternative hypotheses

1. Equal versus not equal hypothesis (two-tailed test)

H_0: parameter = some value

H_1: parameter ≠ some value

2. Equal versus less than (left-tailed test)

H_0: parameter = some value

H_1: parameter < some value

3. Equal versus greater than (right-tailed test)

H_0: parameter = some value

H_1: parameter > some value

The null hypothesis is a statement of ***status quo*** or ***no difference*** and always contains a statement of equality. The null hypothesis is assumed to be true until we have evidence to the contrary. We seek evidence that supports the statement in the alternative hypothesis.

Example

For each of the following claims, determine the null and alternative hypotheses. State whether the test is two-tailed, left-tailed or right-tailed.

a) In 2008, 62% of American adults regularly volunteered their time for charity work. A researcher believes that this percentage is different today.

b) According to a study published in March, 2006 the mean length of a phone call on a cellular telephone was 3.25 minutes. A researcher believes that the mean length of a call has increased since then.

c) Using an old manufacturing process, the standard deviation of the amount of wine put in a bottle was 0.23 ounces. With new equipment, the quality control manager believes the standard deviation has decreased.

Solution

a) In 2008, 62% of American adults regularly volunteered their time for charity work. A researcher believes that this percentage is different today.
The hypothesis deals with a population proportion, p. If the percentage participating in charity work is no different than in 2008, it will be 0.62 so the *null hypothesis is* $H_0 : p = 0.62$.

Since the researcher believes that the percentage is different today, the *alternative hypothesis is a two-tailed hypothesis:* $H_1 : p \neq 0.62$.

b) According to a study published in March, 2006 the mean length of a phone call on a cellular telephone was 3.25 minutes. A researcher believes that the mean length of a call has increased since then.
The hypothesis deals with a population mean, μ. If the mean call length on a cellular phone is no different than in 2006, it will be 3.25 minutes so the *null hypothesis is* $H_0 : \mu = 3.25$

Since the researcher believes that the mean call length has increased, the *alternative hypothesis is:* $H_1 : \mu > 3.25$, *a right-tailed test.*

c) Using an old manufacturing process, the standard deviation of the amount of wine put in a bottle was 0.23 ounces. With new equipment, the quality control manager believes the standard deviation has decreased.
The hypothesis deals with a population standard deviation, σ. If the standard deviation with the new equipment has not changed, it will be 0.23 ounces so the *null hypothesis is* $H_0 : \sigma = 0.23$.

Since the quality control manager believes that the standard deviation has decreased, the *alternative hypothesis is:* $H_1 : \sigma < 0.23$, *a left-tailed test.*

Test Statistic

Definition

The test statistic is a value used in making a decision about the null hypothesis, and is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Test Statistic - Formulas

Test statistic for ***proportion*** $$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Test statistic for ***mean*** $$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad or \quad z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Test statistic for ***standard deviation*** $$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Example

Let's again consider the claim that the XSORT method of gender selection increases the likelihood of having a baby girl. Preliminary results from a test of the XSORT method of gender selection involved 14 couples who gave birth to 13 girls and 1 boy. Use the given claim and the preliminary results to calculate the value of the test statistic. Use the format of the test statistic given above, so that a normal distribution is used to approximate a binomial distribution. (There are other exact methods that do not use the normal approximation.)

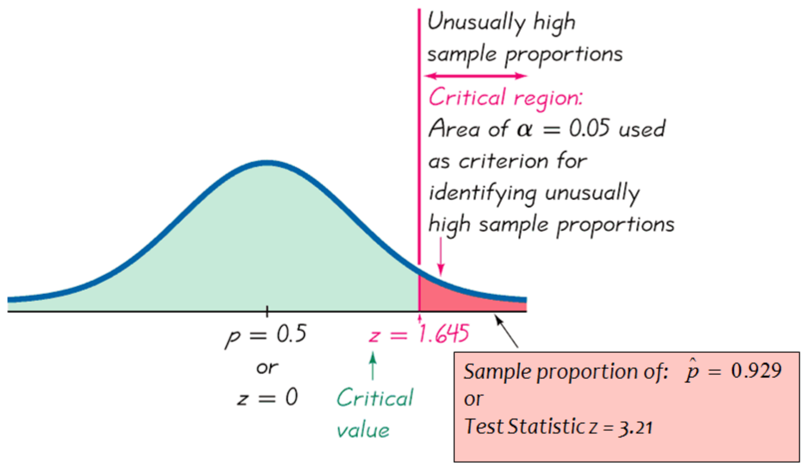
Solution

The claim that the XSORT method of gender selection increases the likelihood of having a baby girl results in the following null and alternative hypothesis $H_0: p = 0.5$ and $H_1: p > 0.5$. We work under the assumption that the null hypothesis is true with $p = 0.5$. The sample proportion of 13 girls in 14 births results in $\hat{p} = \frac{13}{14} = 0.929$. Using $p = 0.5$, $\hat{p} = 0.929$ and $n = 14$, we find the value of the test statistic as follows:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$= \frac{0.929 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{14}}}$$

$$= 3.21$$



We know that a z score of 3.21 *is "unusual" (because it is greater than* 2*). It appears that in addition to being greater than* 0.5*, the sample proportion of* 13/14 *or* 0.929 *is significantly greater than* 0.5*. The figure on the next slide shows that the sample proportion of* 0.929 *does fall within the range of values considered to be significant because they are so far above* 0.5 *that they are not likely to occur by chance (assuming that the population proportion is p* = 0.5).

Critical Region: The ***critical region*** (or ***rejection region***) is the set of all values of the test statistic that cause us to reject the null hypothesis. For example, see the red-shaded region in the previous figure.

Significance Level: The significance level (denoted by α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true. Common choices for α are 0.05, 0.01, and 0.10.

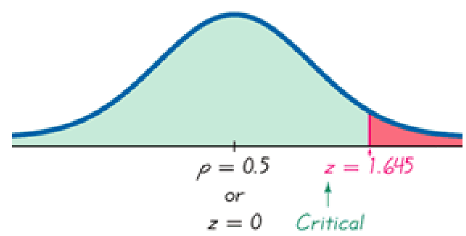
Critical Value: A critical value is any value that separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis. The critical values depend on the nature of the null hypothesis, the sampling distribution that applies, and the significance level α. The critical value of $z = 1.645$ corresponds to a significance level of $\alpha = 0.05$.

Example

Using a significant level of $\alpha = 0.05$, find the critical z value for the alternative hypothesis $H_1 : p > 0.5$ (assuming that the normal distribution can be used to approximate the binomial distribution). This alternative hypothesis is used to test the claim that the XSORT method of gender selection is effective, so that baby girls are more likely, with a proportion greater than 0.5

Solution

With $H_1 : p > 0.5$, the critical region is in the right tail. With the right tail area of 0.05, the critical value is found to be $z = 1.645$. If the right-tailed critical region is 0.05, the cumulative area to the left of the critical value is 0.95 is $z = 1.645$.



Example

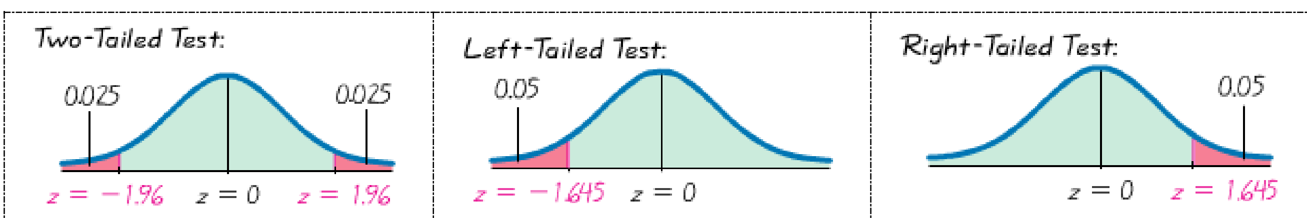
Using a significant level of $\alpha = 0.05$, find the critical z value for the alternative hypothesis $H_1 : p \neq 0.5$ (assuming that the normal distribution can be used to approximate the binomial distribution).

Solution

With $H_1 : p \neq 0.5$, the significant level is 0.05, each of the two tails has an area of 0.025.

From the Normal Distribution Table, $z = -1.96$ and $z = 1.96$ (right side)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233



P-Value

The ***P*-value** (or ***p*-value** or ***probability value***) is the probability of getting a value of the test statistic that is ***at least as extreme*** as the one representing the sample data, assuming that the null hypothesis is true.

Critical region in the ***left*** tail:	*P*-value = area to the ***left*** of the test statistic
Critical region in the ***right*** tail:	*P*-value = area to the ***right*** of the test statistic
Critical region in ***two*** tails:	*P*-value = ***twice*** the area in the tail beyond the test statistic

The null hypothesis is rejected if the *P*-value is very small, such as 0.05 or less.
Here is a memory tool useful for interpreting the *P*-value:

- ✓ If the *P* is low, the null must go.
- ✓ If the *P* is high, the null will fly.

Caution

Don't confuse a *P*-value with a proportion *p*. Know this distinction:

P-value = probability of getting a test statistic at least as extreme as the one representing sample data

p = population proportion

Example

Consider the claim that with the XSORT method of gender selection, the likelihood of having a baby girl is different from $p = 0.5$, and use the test statistic $z = 3.21$ found from 13 girls in 14 births. First determine whether the given conditions result in a critical region in the right tail, left tail, or two tails, then use Figure 8-5 to find the *P*-value. Interpret the *P*-value.

Solution

The claim that the likelihood of having a baby girl is different from $p = 0.5$ can be expressed as $p \neq 0.5$ so the critical region is in two tails. Using Figure 8-5 to find the *P*-value for a two-tailed test, we see that the *P*-value is *twice* the area to the right of the test statistic $z = 3.21$. From Normal Distribution Table, the area to the right of $z = 3.21$ is 0.0007.

In this case, the *P*-value is twice the area to the right of the test statistic, so we have:

$$P\text{–value} = 2 \times 0.0007 = 0.0014$$

- The *P*-value is 0.0014 (or 0.0013 if greater precision is used for the calculations). The small *P*-value of 0.0014 shows that there is a very small chance of getting the sample results that led to a test statistic of $z = 3.21$. This suggests that with the XSORT method of gender selection, the likelihood of having a baby girl is different from 0.5.

Types of Hypothesis Tests: Two-tailed, Left-tailed, Right-tailed

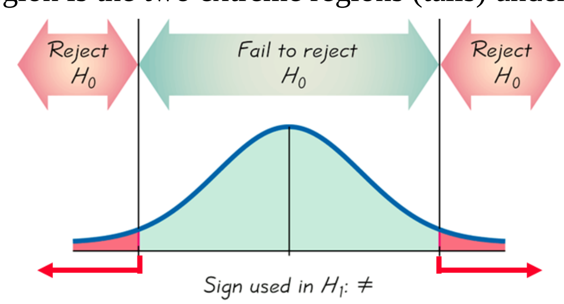
The ***tails*** in a distribution are the extreme regions bounded by critical values.

Determinations of *P*-values and critical values are affected by whether a critical region is in two tails, the left tail, or the right tail. It therefore becomes important to correctly characterize a hypothesis test as two-tailed, left-tailed, or right-tailed.

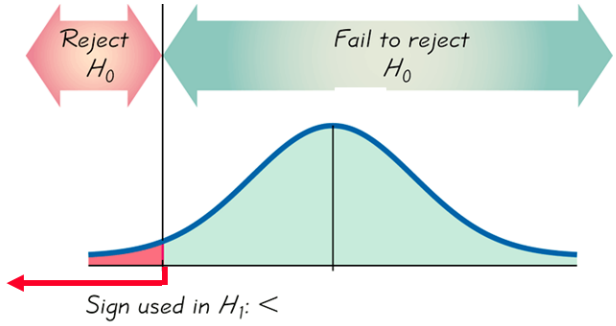
$$H_0 : \; = \quad \& \quad H_1 : \; \neq$$

α is divided equally between the two tails of the critical region

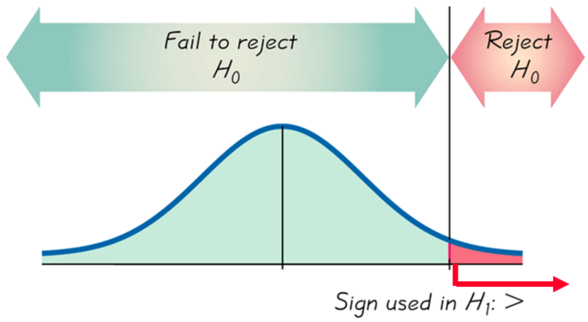
Two–tailed test: The critical region is the two extreme regions (tails) under the curve



Left–tailed test: The critical region is in the extreme region (tail) under the curve



Right–tailed test: The critical region is in the extreme right region (tail) under the curve



Conclusions in Hypothesis Testing

We always test the null hypothesis. The initial conclusion will always be one of the following:

1. Reject the null hypothesis.
2. Fail to reject the null hypothesis.

Decision Criterion

P-value method:

Using the significance level α:

If P-value $\leq \alpha$, ***reject*** H_0.

If P-value $> \alpha$, ***fail to reject*** H_0.

Traditional method:

If the test statistic falls within the critical region, ***reject*** H_0.

If the test statistic does not fall within the critical region, ***fail to reject*** H_0.

Another option:

Instead of using a significance level such as 0.05, simply identify the P-value and leave the decision to the reader.

Confidence Intervals:

A confidence interval estimate of a population parameter contains the likely values of that parameter. If a confidence interval does not include a claimed value of a population parameter, reject that claim.

Example

Suppose a geneticist claims that the XSORT method of gender selection increases the likelihood of a baby girl. This claim of $p > 0.5$ becomes the alternative hypothesis, while the null hypothesis becomes $p = 0.5$. Further suppose that the sample evidence causes us to reject the null hypothesis of $p = 0.5$. State the conclusion in simple, nontechnical terms.

Solution

Because the original claim does not contain equality, it becomes the alternative hypothesis. Because we reject the null hypothesis, we conclude "There is sufficient evidence to support the claim that the XSORT method of gender selection increases the likelihood of a baby girl.

Type I and Type II Errors

Type I error: The mistake of rejecting the null hypothesis when it is actually true. The symbol α (alpha) is used to represent the probability of a type I error.

Type II error: The mistake of failing to reject the null hypothesis when it is actually false. The symbol β (beta) is used to represent the probability of a type II error.

		Reality	
		H_0 is True	H_1 is True
Conclusion	Do Not (Fail to) Reject H_0	Correct Conclusion	Type **II** Error P(type II error) = β
	Reject H_0	Type **I** Error P(type I error) = α	Correct Conclusion

Four Outcomes from Hypothesis Testing

1. Reject the null hypothesis when the alternative hypothesis is true. This decision would be correct.
2. Do not reject the null hypothesis when the null hypothesis is true. This decision would be correct.
3. Reject the null hypothesis when the null hypothesis is true. This decision would be incorrect. This type of error is called a Type I error.
4. Do not reject the null hypothesis when the alternative hypothesis is true. This decision would be incorrect. This type of error is called a Type II error.

Example

Assume that we are conducting a hypothesis test of the claim that a method of gender selection increases the likelihood of a baby girl, so that the probability of a baby girls is $p > 0.5$. Here are the null and alternative hypotheses: $H_0: p = 0.5$, and $H_1: p > 0.5$.

a) Identify a type I error.
b) Identify a type II error.

Solution

a) A type I error is the mistake of rejecting a true null hypothesis, so this is a type I error: Conclude that there is sufficient evidence to support $p > 0.5$, when in reality $p = 0.5$.
b) A type II error is the mistake of failing to reject the null hypothesis when it is false, so this is a type II error: Fail to reject $p = 0.5$ (and therefore fail to support $p > 0.5$) when in reality $p > 0.5$.

Example

For each of the following claims, explain what it would mean to make a Type I error. What would it mean to make a Type II error?

a) In 2008, 62% of American adults regularly volunteered their time for charity work. A researcher believes that this percentage is different today.

b) According to a study published in March, 2006 the mean length of a phone call on a cellular telephone was 3.25 minutes. A researcher believes that the mean length of a call has increased since then.

Solution

a) In 2008, 62% of American adults regularly volunteered their time for charity work. A researcher believes that this percentage is different today.
A ***Type I error*** is made if the researcher concludes that $p \neq 0.62$ when the true proportion of Americans 18 years or older who participated in some form of charity work is currently 62%.
A ***Type II error*** is made if the sample evidence leads the researcher to believe that the current percentage of Americans 18 years or older who participated in some form of charity work is still 62% when, in fact, this percentage differs from 62%.

b) According to a study published in March, 2006 the mean length of a phone call on a cellular telephone was 3.25 minutes. A researcher believes that the mean length of a call has increased since then.
A ***Type I error*** occurs if the sample evidence leads the researcher to conclude that $\mu > 3.25$ when, in fact, the actual mean call length on a cellular phone is still 3.25 minutes.
A ***Type II error*** occurs if the researcher fails to reject the hypothesis that the mean length of a phone call on a cellular phone is 3.25 minutes when, in fact, it is longer than 3.25 minutes.

Controlling Type I and Type II Errors

- For any fixed *a*, an increase in the sample size *n* will cause a decrease in *b*.
- For any fixed sample size *n*, a decrease in *a* will cause an increase in *b*. Conversely, an increase in *a* will cause a decrease in *b*.
- To decrease both *a* and *b*, increase the sample size.

$$\begin{aligned} \alpha &= P\text{ (Type I Error)} \\ &= P\text{ (rejecting } H_0 \text{ when } H_0 \text{ is true)} \end{aligned}$$

$$\begin{aligned} B &= P\text{ (Type II Error)} \\ &= P\text{ (not rejecting } H_0 \text{ when } H_1 \text{ is true)} \end{aligned}$$

The probability of making a Type I error, α, is chosen by the researcher *before* the sample data is collected.
The level of significance, α, is the probability of making a Type I error.

Comprehensive Hypothesis Test

A confidence interval estimate of a population parameter contains the likely values of that parameter. We should therefore reject a claim that the population parameter has a value that is not included in the confidence interval.

Table 8-2 Confidence Level for Confidence Interval

		Two-Tailed Test	One-Tailed Test
Significance Level for Hypothesis Test	0.01	99%	98%
	0.05	95%	90%
	0.10	90%	80%

Definition

The ***power of a hypothesis test*** is the probability ($1 - \beta$) of rejecting a false null hypothesis. The value of the power is computed by using a particular significance level α and a particular value of the population parameter that is an alternative to the value assumed true in the null hypothesis.
That is, the power of the hypothesis test is the probability of supporting an alternative hypothesis that is true.

Power and the Design of Experiments

Just as 0.05 is a common choice for a significance level, a power of at least 0.80 is a common requirement for determining that a hypothesis test is effective. (Some statisticians argue that the power should be higher, such as 0.85 or 0.90.) When designing an experiment, we might consider how much of a difference between the claimed value of a parameter and its true value is an important amount of difference. When designing an experiment, a goal of having a power value of at least 0.80 can often be used to determine the minimum required sample size.

Exercises Section 3.5 – Language of Hypothesis Testing

1. Bottles of Bayer aspirin are labeled with a statement that the tablets each contain 325 mg of aspirin. A quality control manager claims that a large sample of data can be used to support the claim that the mean amount of aspirin in the tablets is equal to 325 mg, as the label indicates. Can a hypothesis test be used to support that claim? Why or Why not?

2. In the preliminary results from couples using the Gender Choice method of gender selection to increase the likelihood of having a baby girl, 20 couples used the Gender Choice method with the result that 8 of them had baby girls and 12 had baby boys. Given that the sample proportion of girls is $\frac{8}{20}$ *or* 0.4, can the sample data support the claim that the proportion of girls is greater than 0.5? Can any sample proportion less than 0.5 be used to support a claim that the population proportion is greater than 0.5?

3. Express the null hypothesis H_0 and alternative hypothesis H_1 in symbolic form. Be sure to use the correct symbol (μ, p, σ) for indicated parameter

a) The mean annual income of employees who took a statistics course is greater than $60,000.
b) The proportion of people aged 18 to 25 who currently use illicit drugs is equal to 0.20 (or 20%).
c) The standard deviation of human body temperatures is equal to 0.62°F.
d) The majority of college students have credit cards.
e) The proportion of homes with fire extinguishers is 0.80.
f) The mean weight of plastic discarded by households in one week is less than 1 kg.

4. Assume that the normal distribution applies and find the critical *z* values.

a) Two-tailed test: $\alpha = 0.01$.
b) Right-tailed test: $\alpha = 0.02$.
c) Left-tailed test: $\alpha = 0.10$.
d) $\alpha = 0.05;\quad H_1 \text{ } is \text{ } p \neq 0.4$
e) $\alpha = 0.01;\quad H_1 \text{ } is \text{ } p > 0.5$
f) $\alpha = 0.005;\quad H_1 \text{ } is \text{ } p < 0.8$
g) $\alpha = 0.05$ for two-tailed test
h) $\alpha = 0.05$ for left-tailed test
i) $\alpha = 0.08;\quad H_1 \text{ } is \text{ } \mu \neq 3.25$

5. The claim is that the proportion of peas with yellow pods is equal to 0.25 (or 25%). The sample statistics from one of Mendel's experiments include 580 peas with 152 of them having yellow pods. Find the value of the test statistic *z* using $z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$

6. The claim is that less than $\frac{1}{2}$ of adults in U.S. have carbon monoxide detectors. A KRC Research survey of 1005 adults resulted in 462 who have carbon monoxide detectors. Find the value of the test statistic z using $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

7. The claim is that more than 25% of adults prefer Italian food as their favorite ethnic food. A Harris Interactive survey of 1122 adults resulted in 314 who say that Italian food is their favorite ethnic food. Find the value of the test statistic z using $z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$

8. Find P-value by using a 0.05 significance level and state the conclusion about the null hypothesis. (reject the null hypothesis or fail to reject the null hypothesis)

a) The test statistic in a left-tailed test is $z = -1.25$

b) The test statistic in a right-tailed test is $z = 2.50$

c) The test statistic in a two-tailed test is $z = 1.75$

d) With $H_1 : p \neq 0.707$, the test statistic is $z = -2.75$

e) With $H_1 : p > \frac{1}{4}$, the test statistic is $z = 2.30$

f) With $H_1 : p < 0.777$, the test statistic is $z = -2.95$

9. The percentage of nonsmokers exposed to secondhand smoke is equal to 41%. Identify the type I error and type II error.

10. The percentage of Americans who believe that life exists only on earth is equal to 20%. Identify the type I error and type II error.

11. The percentage of college students who consume alcohol is greater than 70%. Identify the type I error and type II error.

12. An entomologist writes an article in a scientific journal which claims that fewer than 13 in 10,000 male fireflies are unable to produce light due to a genetic mutation. Use the parameter p, the true proportion of fireflies unable to produce light. Express the null hypothesis and the alternative hypothesis in symbolic form. $(\mu,\ p,\ \sigma)$

Section 3.6 – Hypothesis Tests for a Population Proportion

A researcher obtains a random sample of 1000 people and finds that 534 are in favor of the banning cell phone use while driving, so $\hat{p}$ = 534/1000. Does this suggest that more than 50% of people favor the policy? Or is it possible that the true proportion of registered voters who favor the policy is some proportion less than 0.5 and we just happened to survey a majority in favor of the policy? In other words, would it be unusual to obtain a sample proportion of 0.534 or higher from a population whose proportion is 0.5? What is convincing, or statistically significant, evidence?

When observed results are unlikely under the assumption that the null hypothesis is true, we say the result is ***statistically significant***. When results are found to be statistically significant, we reject the null hypothesis.

Basic Methods of Testing Claims about a Population Proportion *p*

Notation

n = number of trials

$\hat{p} = \frac{x}{n}$ (***sample*** proportion)

p = population proportion (used in the null hypothesis)

$q = 1 - p$

Obtaining $\hat{p}$

$\hat{p}$ sometimes is given directly "10% of the observed sports cars are red" is expressed as $\hat{p} = 0.10$

$\hat{p}$ sometimes must be calculated "96 surveyed households have cable TV and 54 do not" is calculated using

$$\hat{p} = \frac{x}{n} = \frac{96}{96+54} = 0.64$$

Requirements for Testing Claims about a Population Proportion *p*

1. The sample observations are a simple random sample.
2. The conditions for a ***binomial distribution*** are satisfied.
3. The conditions $np \geq 5$ and $nq \geq 5$ are both satisfied, so the binomial distribution of sample proportions can be approximated by a normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$. Note: p is the assumed proportion not the sample proportion.

Test Statistic for Testing a Claim about a Proportion

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

***P*-values:** Use the standard Normal Distribution Table and refer to Figure

Critical Values: Use the standard Normal Distribution Table

Caution: Don't confuse a *P*-value with a proportion *p*.

P-value = probability of getting a test statistic at least as extreme as the one representing sample data
p = population proportion

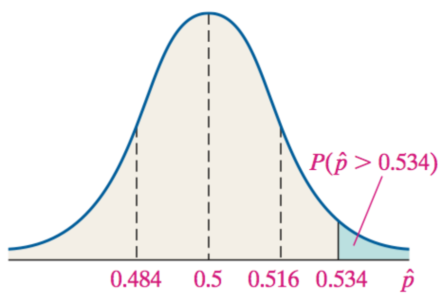
- When testing claims about a population proportion, the traditional method and the *P*-value method are equivalent and will yield the same result since they use the same standard deviation based on the ***claimed proportion*** *p*. However, the confidence interval uses an estimated standard deviation based upon the ***sample proportion*** $\hat{p}$. Consequently, it is possible that the traditional and *P*-value methods may yield a different conclusion than the confidence interval method.
  A good strategy is to use a confidence interval to estimate a population proportion, but use the *P*-value or traditional method for testing a claim about the proportion.

P–Value Method

If the sample proportion of getting a sample proportion as extreme or more extreme than the one obtained is small under the assumption the statement in the null hypothesis is true, reject the null hypothesis.

The Logic of the *P*-Value Approach

A second criterion we may use for testing hypotheses is to determine how likely it is to obtain a sample proportion of $\hat{p}$ = 534/1000 = 0.534 or higher from a population whose proportion is 0.5. If a sample proportion of 0.534 or higher is unlikely (or unusual), we have evidence against the statement in the null hypothesis. Otherwise, we do not have sufficient evidence against the statement in the null hypothesis.



We can compute the probability of obtaining a sample proportion of 0.534 or higher from a population whose proportion is 0.5 using the normal model

Example

A sample proportion of 0.534 (random sample of 1000 people out 534) or higher from a population whose proportion is 0.5.If a sample proportion of 0.534 or higher is unlikely (or unusual), Find *P*-value.

Solution

$$z = \frac{0.534 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{1000}}} = 2.15$$

$$P(\hat{p} > 0.534) = P(z > 2.15) = 1 - 0.9842 \underline{= 0.0158}$$

The value 0.0158 is called the *P*-value, which means about 2 samples in 100 will give a sample proportion as high or higher than the one we obtained if the population proportion really is 0.5. Because these results are unusual, we take this as evidence against the statement in the null hypothesis.

Example

The text refers to a study in which 57 out of 104 pregnant women correctly guessed the sex of their babies. Use these sample data to test the claim that the success rate of such guesses is no different from the 50% success rate expected with random chance guesses. Use a 0.05 significance level.

Solution

Requirements are satisfied: simple random sample; fixed number of trials (104) with two categories (guess correctly or do not)

$np = (104)(0.5) = 52 \geq 5$

$nq = (104)(0.5) = 52 \geq 5$

Step 1: Original claim is that the success rate is no different from 50%: $p = 0.50$

Step 2: Opposite of original claim is $p \neq 0.50$

Step 3: $p \neq 0.50$ does not contain equality so it is H_1.

$H_0 : p = 0.50$ null hypothesis and original claim

$H_1 : p \neq 0.50$ alternative hypothesis

Step 4: significance level is $\alpha = 0.50$

Step 5: sample involves proportion so the relevant statistic is the sample proportion, $\hat{p}$

Step 6: calculate z:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{\frac{57}{104} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{104}}}$$

$$= 0.98$$

Two-tailed test, P-value is twice the area to the right of test statistic

From the Normal Distribution Table; $z = 0.98$ has an area of 0.8365 to its left, so area to the right is $1 - 0.8365 = 0.1635$, doubles yields 0.3270 (technology provides a more accurate P–value of 0.3268

Example

Suppose a geneticist claims that the XSORT method of gender selection. Among 726 babies born to couples using the XSORT method in an attempt to have a baby girl, 668 of the babies were girls and the others were boys. Use these results with a 0.05 significant level to test the claim that among babies born to couples using the XSORT method, the proportion of girls is greater than the value of 0.5 that is expected with no treatment. Here is a summary of the claim and the sample data:

Solution

Claim: With the XSORT method, the proportion of girls $p > 0.5$

Sample data: $n = 726$ *and* $\hat{p} = \frac{668}{726} = 0.920$

***Step* 1:** The original claim is symbolic is $p > 0.5$

***Step* 2:** The opposite of the original claim is $p \le 0.5$

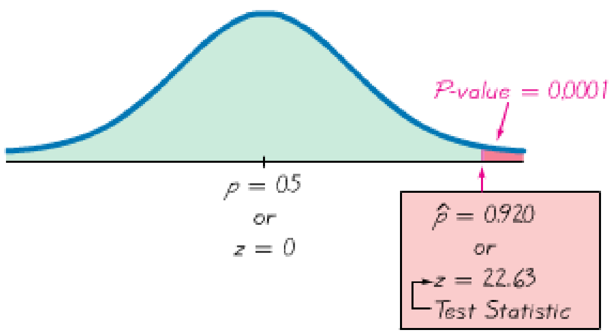
***Step* 3:** $p > 0.5$ does not contain equality, so it becomes the alternative hypothesis. The null hypothesis is the statement that *p* equals the fixed value of 0.5.

$H_0 : p = 0.5$

$H_1 : p > 0.5$

***Step* 4:** The significant level of $\alpha = 0.05$, which is a very common choice.

***Step* 5:** Because we are testing a claim about a population proportion *p*, the sample statistic $\hat{p}$ is relevant to this test. The sampling distribution of sample proportion $\hat{p}$ can be approximated by a normal distribution.



***Step* 6:** The test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} = \frac{0.920 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{726}}} = 22.63$$

P–values are:

Left–tailed test: *P*–value = area to left of test statistic *z*

Right–tailed test: *P*–value = area to right of test statistic *z*

Two–tailed test: *P*–value = twice the area of the extreme region bounded by test statistic *z*

Because the hypothesis test we are considering is right-tailed with a test statistic of $z = 22.63$, the *P*-value is the area to the right of $z = 22.63$. Referring to Normal Distribution Table, for values of $z = 3.50$ and higher, we use 0.0001 for the cumulative area to the *right* of the test statistic. The *P*-value is therefore 0.0001.

***Step* 7:** Because the *P*-value is 0.0001 is less than or equal to the significance level of $\alpha = 0.05$, we reject the null hypothesis

***Step* 8:** **W**e conclude that there is sufficient sample evidence to support the claim that among babies born to couples using the XSORT method, the proportion of girls is greater than 0.5.

Testing Hypotheses Claims

To test hypotheses regarding the population proportion, we can use the steps that follow, provided that:

- ✓ The sample is obtained by simple random sampling.
- ✓ $np_0\left(1-p_0\right) \geq 0$
- ✓ The sampled values are independent of each other.

***Step* 1**: Determine the null and alternative hypotheses. The hypotheses can be structured in one of three ways:

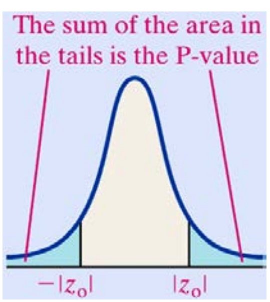
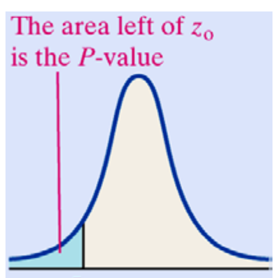
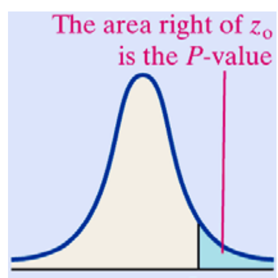
Two–Tailed	***Left–Tailed***	***Right–Tailed***
$H_0 : p = p_0$	$H_0 : p = p_0$	$H_0 : p = p_0$
$H_1 : p \neq p_0$	$H_1 : p < p_0$	$H_1 : p > p_0$

p_0 is assumed value of the population proportion.

***Step* 2:** Select a level of significance, α, based on the seriousness of making a Type I error.

Step* 3**: Compute the ***test statistic $z_0 = \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0\left(1-p_0\right)}{n}}}$

***Step* 4:** Compare the critical value with the test statistic:

Two–Tailed	***Left–Tailed***	***Right–Tailed***
$z_0 < -z_{\alpha/2}$ *or* $z_0 > z_{\alpha/2}$ *Reject the null hypothesis*	$z_0 < -z_{\alpha}$ *Reject the null hypothesis*	$z_0 > z_{\alpha}$ *Reject the null hypothesis*
		

***Step* 5:** If the *P*-value < α, reject the null hypothesis.

***Step* 6:** State the conclusion.

Exercises *Section* 3.6 – Hypothesis Tests for a Population Proportion

1. In a Harris poll, adults were asked if they are in favor of abolishing the penny. Among the responses, 1261 answered "no", and 491 answered "yes", and 384 had no opinion. What is the sample proportion of ***yes*** responses, and what notation is used to represent it?

2. A recent study showed that 53% of college applications were submitted online. Assume that this result is based on a simple random sample of 1000 college applications, with 530 submitted online. Use a 0.01 significance level to test the claim that among all college applications the percentage submitted online is equal to 50%

a) What is the test statistic?
b) What are the critical values?
c) What is the *P*-Value?
d) What is the conclusion?
e) Can a hypothesis test be used to "prove" that the percentage of college applications submitted online is equal to 50% as claimed?

3. In a survey, 1864 out of 2246 randomly selected adults in the U.S. said that texting while driving should be illegal. Consider a hypothesis test that uses a 0.05 significance level to test the claim that more than 80% of adults believe that testing while driving should be illegal

a) What is the test statistic?
b) What are the critical values?
c) What is the *P*-Value?
d) What is the conclusion?

4. In a Pew Research Center poll of 745 randomly selected adults, 589 said that it is morally wrong to not report all income on tax returns. Use a 0.01 significance level to test the claim that 75% of adults say that it is morally wrong to not report all income on tax returns.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

5. 308 out of 611 voters surveyed said that they voted for the candidate who won. Use a 0.01 significance level to test the claim that among all voters, the percentage who believe that they voted for the winning candidate is equal to 43%, which is the actual percentage of votes for the winning candidate. What does the result suggest about voter perceptions?
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

6. The company Drug Test Success provides a "1-Panel-THC" test for marijuana usage. Among 300 tested subjects, results from 27 subjects were wrong (either a false positive or a false negative). Use a 0.05 significance level to test the claim that less than 10% of the test results are wrong. Does the test appear to be good for most purposes?

7. When testing gas pumps in Michigan for accuracy, fuel-quality enforcement specialists tested pumps and found that 1299 of them were not pumping accurately (within 3.3 oz. when 5 gal. is pumped), and 5686 pumps were accurate. Use a 0.01 significance level to test the claim of an industry representative that less than 20% of Michigan gas pumps are inaccurate. From the perspective of the consumer, does that rate appear to be low enough?

8. Trials in an experiment with a polygraph include 98 results that include 24 cases of wrong results and 74 cases of correct results. Use a 0.05 significance level to test the claim that such polygraph results are correct less than 80% of the time. Based on the results, should polygraph test results be prohibited as evidence in trials?

9. In recent years, the Town of Newport experienced an arrest rate of 25% for robberies. The new sheriff compiles records showing that among 30 recent robberies, the arrest rate is 30%, so she claims that her arrest rate is greater than the 25% rate in the past. Is there sufficient evidence to support her claim that the arrest rate is greater than 25%?

10. A survey showed that among 785 randomly selected subjects who completed 4 years of college, 18.3 % smoke and 81.7% do not smoke. Use a 0.01 significance level to test the claim that the rate of smoking among those with 4 years of college is less than the 27% rate for the general population. Why would college graduates smoke at a lower rate than others?

11. When 3011 adults were surveyed, 73% said that they use the Internet. Is it okay for a newspaper reporter to write that "3/4 of all adults use the internet"? Why or Why not?

12. A hypothesis test is performed to test the claim that a population proportion is greater than 0.7. Find the probability of a type II error, β, given that the true value of the population proportion is 0.72. The sample size is 50 and the significance level is 0.05.

13. In a sample of 88 children selected randomly from one town, it is found that 8 of them suffer asthma. Find the *P*-value for a test of the claim that the proportion of all children in the town who suffer from asthma is equal to 11%.

14. An airline claims that the no-show rate for passengers booked on its flights is less than 6%. Of 380 randomly selected reservation, 18 were no-shows. Find the P-value for a test of the airline's claim.

15. In 1997, 46% of Americans said they did not trust the media "when it comes to reporting the news fully, accurately and fairly". In a 2007 poll of 1010 adults nationwide, 525 stated they did not trust the media. At the $\alpha = 0.05$ level of significance, is there evidence to support the claim that the percentage of Americans that do not trust the media to report fully and accurately has increased since 1997?

16. In 2006, 10.5% of all live births in the United States were to mothers under 20 years of age. A sociologist claims that births to mothers under 20 years of age is decreasing. She conducts a simple random sample of 34 births and finds that 3 of them were to mothers under 20 years of age. Test the sociologist's claim at the $\alpha = 0.01$ level of significance.

Section 3.7 – Hypothesis Tests for a Population Mean

Objective

Test a claim about a population mean (with σ known) by using a formal method of hypothesis testing.

Notation

n = Sample size
$\overline{x}$ = sample mean
$\mu_{\overline{x}}$ = population mean of all sample means from samples of size n
σ = known value of the population standard deviation

Requirements for Testing Claims about a Population Mean (with σ Known)

1. The sample is a simple random sample.
2. The value of the population standard deviation σ is known.
3. Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$.

Test Statistic for Testing a Claim About a Mean (with σ Known)

To test hypotheses regarding the population mean assuming the population standard deviation is unknown, we use the t-distribution rather than the Z–distribution. When we replace σ with s

$$z = \frac{\overline{x} - \mu_{\overline{x}}}{\frac{\sigma}{\sqrt{n}}}$$

follows **Student's *t*-distribution** with $n - 1$ degrees of freedom.

Properties of the *t*–Distribution

1. The t-distribution is different for different degrees of freedom.
2. The t-distribution is centered at 0 and is symmetric about 0.
3. The area under the curve is 1. Because of the symmetry, the area under the curve to the right of 0 equals the area under the curve to the left of 0 equals 1/2.
4. As t increases (or decreases) without bound, the graph approaches, but never equals, 0.
5. The area in the tails of the t-distribution is a little greater than the area in the tails of the standard normal distribution because using s as an estimate of σ introduces more variability to the t-statistic.
6. As the sample size n increases, the density curve of t gets closer to the standard normal density curve. This result occurs because as the sample size increases, the values of s get closer to the values of σ by the Law of Large Numbers.

Testing Hypotheses Regarding a Population Mean

To test hypotheses regarding the population mean, we use the following steps, provided that:

- ✓ The sample is obtained using simple random sampling.
- ✓ The sample has no outliers, and the population from which the sample is drawn is normally distributed or the sample size is large ($n \geq 30$).
- ✓ The sampled values are independent of each other.

***Step* 1**: Determine the null and alternative hypotheses. The hypotheses can be structured in one of three ways:

Two–Tailed	***Left–Tailed***	***Right–Tailed***
$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$

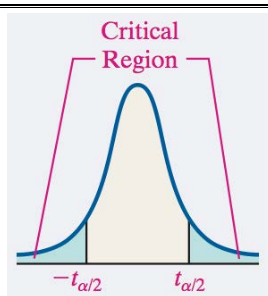
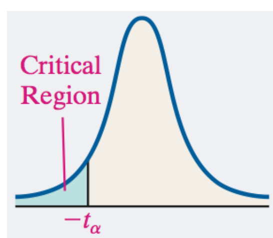
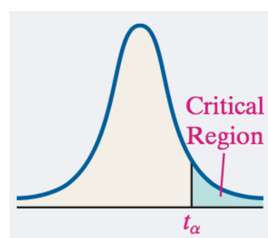
μ_0 is assumed value of the population mean.

***Step* 2:** Select a level of significance, α, based on the seriousness of making a Type I error.

Classical Approach

Step* 3**: Compute the ***test statistic $t_0 = \dfrac{\overline{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

which follows the Student's t-distribution with $n - 1$ degrees of freedom.

Two–Tailed	***Left–Tailed***	***Right–Tailed***
		

P-Value Approach

Step* 3:** Compute the ***test statistic:

***Step* 4:** If the P-value $< \alpha$, reject the null hypothesis.

***Step* 5:** State the conclusion.

The procedure is robust, which means that minor departures from normality will not adversely affect the results of the test. However, for small samples, if the data have outliers, the procedure should not be used.

Example

People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: n = 40 and $\bar{x}$ = 172.55 lb. Research from several other sources suggests that the population of weights of men has a standard deviation given by σ = 26 lb. Use these results to test the claim that men have a mean weight greater than 166.3 lb, which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and use the *P*-value method outlined in Figure 8-8.

Solution

Requirements are satisfied: simple random sample, σ is known (26 lb), sample size is 40 (n > 30)

Step 1: Express claim as μ > 166.3 lb

Step 2: alternative to claim is $\mu \le$ 166.3 lb

Step 3: μ > 166.3 lb does not contain equality, it is the alternative hypothesis:

H_0 : μ = 166.3 lb. null hypothesis

H_1 : μ > 166.3 lb. alternative hypothesis and original claim

Step 4: significance level is α = 0.50

Step 5: claim is about the population mean, so the relevant statistic is the sample mean (172.55 lb), σ is known (26 lb), sample size greater than 30

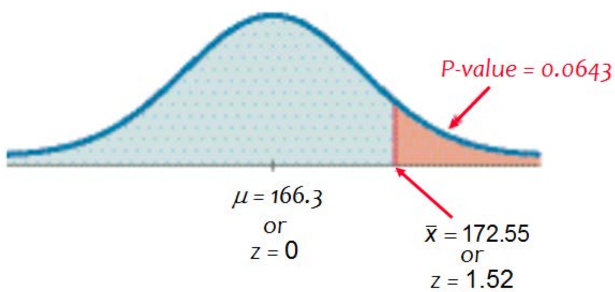
Step 6: calculate z:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26}{\sqrt{40}}} = 1.52$$

Right--tailed test, so *P*-value is the area is to the right of z = 1.52;

From the Normal Distribution Table; area to the left of z = 1.52 is 0.9357, so the area to the right is 1 – 0.9357 = 0.0643. The *P*-value is 0.0643

***Step* 7:** The *P*-value of 0.0643 is greater than the significance level of α = 0.05, we fail to reject the null hypothesis.



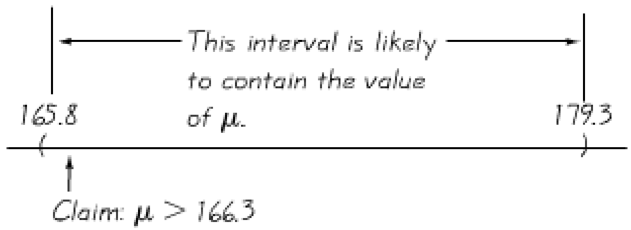
- ✓ The *P*-value of 0.0643 tells us that if men have a mean weight given by μ = 166.3 lb, there is a good chance (0.0643) of getting a sample mean of 172.55 lb. A sample mean such as 172.55 lb could easily occur by chance. There is not sufficient evidence to support a conclusion that the population mean is greater than 166.3 lb, as in the National Transportation and Safety Board's recommendation.

The ***traditional method***: Use z = 1.645 instead of finding the P-value. Since z = 1.52 does not fall in the critical region, again fail to reject the null hypothesis.

Confidence Interval method: Use a one-tailed test with a = 0.05, so construct a 90% confidence interval:

$165.8 < \mu < 179.3$

The confidence interval contains 166.3 lb, we cannot support a claim that μ is greater than 166.3. Again, fail to reject the null hypothesis.



Underlying Rationale of Hypothesis Testing

If, under a given assumption, there is an extremely small probability of getting sample results at least as extreme as the results that were obtained, we conclude that the assumption is probably not correct. When testing a claim, we make an assumption (null hypothesis) of equality. We then compare the assumption and the sample results and we form one of the following conclusions:

- If the sample results (or more extreme results) can easily occur when the assumption (null hypothesis) is true, we attribute the relatively small discrepancy between the assumption and the sample results to chance.
- If the sample results cannot easily occur when that assumption (null hypothesis) is true, we explain the relatively large discrepancy between the assumption and the sample results by concluding that the assumption is not true, so we reject the assumption.

Distinguish between *Statistical* Significance and *Practical* Significance

When a large sample size is used in a hypothesis test, the results could be statistically significant even though the difference between the sample statistic and mean stated in the null hypothesis may have no ***practical significance***.

> ***Practical significance*** refers to the idea that, while small differences between the statistic and parameter stated in the null hypothesis are statistically significant, the difference may not be large enough to cause concern or be considered important.

Large sample sizes can lead to statistically significant results while the difference between the statistic and parameter is not enough to be considered practically significant.

Exercises Section 3.7 – Hypothesis Tests for a Population Mean

1. Because the amounts of nicotine in king size cigarettes listed below

1.1	1.7	1.7	1.1	1.1	1.4	1.1	1.4	1	1.2	1.1	1.1	1.1
1.1	1.1	1.8	1.6	1.1	1.2	1.5	1.3	1.1	1.3	1.1	1.1	

We must satisfy the requirement that the population is normally distributed. How do we verify that a population is normally distributed?

2. If you want to construct a confidence interval to be used for testing the claim that college students have a mean IQ score that is greater than 100, and you want the test conducted with a 0.01 significance level, what confidence level should be used for the confidence interval?

3. A jewelry designer claims that women have wrist breadths with a mean equal to 5 cm. A simple random sample of the wrist breadths of 40 women has a mean of 5.07 cm. Assume that the population standard deviation is 0.33 cm. Use the accompanying TI display to test the designer's claim.

```
Z-Test
 μ≠5
 z=1.341572341
 p=.1797348219
 x̄=5.07
 n=40
```

Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

4. The U.S. Mint has a specification that pennies have a mean weight of 2.5 g. Assume that weights of pennies have a standard deviation of 0.0165 g and use the accompanying Minitab display to test the claim that the sample is from a population with a mean that is less than 2.5 g. These Minitab results were obtained using the 37 weights of post 1983 pennies.

```
Test of mu = 2.5 vs < 2.5. Assumed s.d. = 0.0165
                                95% Upper
 N      Mean      StDev      Bound          Z        P
37   2.49910   0.01648    2.50356     -0.33    0.370
```

5. In the manual "How long to have a Number One the Easy Way," by KLF Publications, it is stated that a song "must be no longer than 3 minutes and 30 seconds" (or 210 seconds). A simple random sample of 40 current hit songs results in a mean length of 252.5 sec. Assume that the standard deviation of song lengths is 54.5 sec. Use a 0.05 significance level to test the claim that the sample is from a population of songs with a mean greater than 210 sec. What do these results suggest about the advice given in the manual?

6. A simple random sample of 50 adults is obtained, and each person's red blood cell count (in cells per microliter) is measured. The sample mean is 5.23. The population standard deviation for red blood cell counts is 0.54. Use a 0.01 significance level to test the claim that the sample is from population with a mean less than 5.4, which is value often used for the upper limit of the range of normal values. What do the results suggest about the sample group?

7. A simple random sample of 106 body temperature with a mean of 98.20 °F. Assume that σ is known to be 0.62 °F. Use a 0.05 significance level to test the claim that the mean body temperature of the population is equal to 98.6 °F, as is commonly believed. Is there sufficient evidence to conclude that the common belief is wrong?

8. When 40 people used the Weight Watchers diet for one year, their mean weight loss was 3.0 lb. Assume that the standard deviation of all such weight changes is $\sigma = 4.9$ lb. and use a 0.01 significance level to test the claim that the mean weight loss is greater than 0. Based on these results, does the diet appear to be effective? Does the diet appear to have a practical significance?

9. The health of the bear population in Yellowstone National Park is monitored by periodic measurements taken from anesthetized bears. A sample of 54 bears has a mean weight of 182.9 lb. Assuming that σ is known to be 121.8 lb. use a 0.05 significance level to test the claim that the population mean of all such bear weights is greater than 150 lb.

10. A simple random sample of 401 salaries of NCAA football coaches in the NCAA has a mean of \$415,953. The standard deviation of all salaries of NCAA football coaches is \$463,364. Use a 0.05 significance level to test the claim that the mean salary of a football coach in the NCAA is less than \$500,000.

11. A simple random sample of 36 cans of regular Coke has a mean volume of 12.19 oz. Assume that the standard deviation of all cans of regular Coke is 0.11 oz. Use a 0.01 significance level to test the claim that cans if regular Coke have volumes with a mean of 12 oz., as stated on the label. If there is a difference, is it substantial?

12. A simple random sample of FICO credit rating scores is obtained, and the scores are listed below.

714	751	664	789	818	779	698	836	753	834	693	802

As the writing, the mean FICO score was reported to be 678. Assuming the standard deviation of all FICO scores is known to be 58.3, use a 0.05 significance level to test the claim that these sample FICO scores come from a population with a mean equal to 678.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

13. Listed below are recorded speeds (in mi/h) of randomly selected cars traveling on a section of Highway 405 in Los Angeles.

68	68	72	73	65	74	73	72	68	65	65	73	66	71	68	74	66	71	65	73
59	75	70	56	66	75	68	75	62	72	60	73	61	75	58	74	60	73	58	75

That part of the highway has posted speed limit of 65 mi/h. Assume that the standard deviation od speeds is 5.7 mi/h and use a 0.01 significance level to test the claim that the sample is from a population with a mean that is greater than 65 *mi/h*.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

14. Given a simple random sample of speeds of cars on Highway in CA, you want to test the claim that the sample that the sample values are from a population with a mean greater than the posted speed limit of 65 *mi/hr*. Is it necessary to determine whether the sample is from a normally distributed population? If so, what methods can be used to make that determination?

15. In statistics, what does ***df*** denote. If a simple random sample of 20 speeds of cars is to be used to test the claim that the sample values are from a population with a mean greater than the posted speed limit of 65 mi/h, what is the specific value of *df*?

16. Claim about IQ scores of statistics instructors: $\mu > 100$, sample data: $n = 15, \quad \bar{x} = 118, \quad s = 11$. The sample data appear to come from a normally distributed population with unknown μ and σ. Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

17. Claim about FICO credit scores of adults: $\mu = 678$, sample data: $n = 12, \quad \bar{x} = 719, \quad s = 92$. The sample data appear to come from a population with a distribution that is not normal, and σ Is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

18. Claim about daily rainfall amounts in Boston: $\mu < 0.20\,in$, sample data: $n = 19, \bar{x} = 0.10\,in, s = 0.26\,in.$

19. The sample data appear to come from a population with a distribution that is very far from normal, and σ is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

20. Testing a claim about the mean weight of M&M's: Right-tailed test with n = 25 and test statistic t = 0.430. Find the *P*-value and find a range of values for the *P*-value.

21. Test a claim about the mean body temperature of healthy adults: left-tailed test with n = 11 and test statistic t = –3.158. Find the *P*-value and find a range of values for the *P*-value.

22. Two-tailed test with n = 15 and test statistic t = 1.495. Find the *P*-value and find a range of values for the *P*-value.

23. In an analysis investigating the usefulness of pennies, the cents portions of 100 randomly selected checks are recorded. The sample has mean of 23.8 *cents* and a standard deviation of 32.0 *cents*. If the amounts from 0 *cents* to 99 *cents* are all equally likely, the mean is expected to be 49.5 *cents*. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 49.5 *cents*. What does the result suggest about the cents portions of the checks?

24. A simple random sample of 40 recorded speeds (in *mi/h*) is obtained from cars traveling on a section of Highway 405 in Los Angeles. The sample has a mean of 68.4 *mi/h* and a standard deviation of 5.7 mi/h. Use a 0.05 significance level to test the claim that the mean speed of all cars is greater than the posted speed limit of 65 *mi/h*.

25. The heights are measured for the simple random sample of supermodels. They have mean height of 70.0 *in*. and a standard deviation of 1.5 *in*. Use a 0.01 significance level to test the claim that supermodels have heights with a mean that is greater than the mean heights of 63.6 *in*. for women in general population. Given that there are only nine heights represented, can we really conclude that supermodels are taller than the typical woman?

26. The National Highway Traffic Safety Administration conducted crash tests of child booster seats for cars. Listed below are results from those tests, with measurement given in hic (standard *head injury condition* units). The safety requirement is that the hic measurement should be less than 1000 *hic*. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 1000 *hic*.

774 649 1210 546 431 612

Do the results suggest that all of the child booster seats meet the specified requirement?

27. The trend of thinner Miss America winners has generated charges that the contest encourages unhealthy diet habits among young women. Listed below are body mass indexes (BMI) for recent Miss America winners. Use a 0.01 significance level to test the claim that recent Miss America winners are from a population with a mean BMI less than 20.16, which was the BMI for winners from the 1920s and 1930s.

19.5 20.3 19.6 20.2 17.8 17.9 19.1 18.8 17.6 16.8

Do recent winners appear to be significantly different from those in the 1920s and 1930s?

28. The list measured voltage amounts supplied directly to the author's home

123.8	123.9	123.9	123.3	123.4	123.3	123.3	123.6	123.5	123.5	123.5	123.7
123.6	123.7	123.9	124.0	124.2	123.9	123.8	123.8	124.0	123.9	123.6	123.5
123.4	123.4	123.4	123.4	123.3	123.3	123.5	123.6	123.8	123.9	123.9	123.8
123.9	123.7	123.8	123.8								

The Central Hudson power supply company states that it has a target power supply of 120 volts. Using those home voltage amounts, test the claim that the mean is 120 volts. Use a 0.01 significance level.

29. When testing a claim about a population mean with a simple random sample selected from a normally distributed population with unknown σ, the student t distribution should be used for finding critical values and/or a *P*–value. If the standard normal distribution is incorrectly used instead, does that mistake make you more or less likely to reject the null hypothesis, or does it not make a difference? Explain.

30. The list measured human body temperature.

98.6	98.6	98.0	98.0	99.0	98.4	98.4	98.4	98.4	98.6	98.6	98.8	98.6	97.0	97.0
98.8	97.6	97.7	98.8	98.0	98.0	98.3	98.5	97.3	98.7	97.4	98.9	98.6	99.5	97.5
97.3	97.6	98.2	99.6	98.7	99.4	98.2	98.0	98.6	98.6	97.2	98.4	98.6	98.2	98.0
97.8	98.0	98.4	98.6	98.6	97.8	99.0	96.5	97.6	98.0	96.9	97.6	97.1	97.9	98.4
97.3	98.0	97.5	97.6	98.2	98.5	98.8	98.7	97.8	98.0	97.1	97.4	99.4	98.4	98.6
98.4	98.5	98.6	98.3	98.7	98.8	99.1	98.6	97.9	98.8	98.0	98.7	98.5	98.9	98.4

98.6	97.1	97.9	98.8	98.7	97.6	98.2	99.2	97.8	98.0	98.4	97.8	98.4	97.4	98.0
97.0														

Use the temperatures listed for 12 AM on day 2 to test the common belief that the mean body temperature is 98.6 °F. Does that common belief appear to be wrong?

31. Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither

a) Claim $\mu = 981$. Sample data: $n = 20, \quad \bar{x} = 946, \quad s = 27$. The sample data appear to come from a normally distributed population with $\sigma = 30$.

b) Claim $\mu = 105$. Sample data: $n = 16, \quad \bar{x} = 101, \quad s = 15.1$. The sample data appear to come from a normally distributed population with unknown μ *and* σ.

32. Assume the resting metabolic rate (RMR) of healthy males in complete silence is 5710 kJ/day. Researchers measured the RMR of 45 healthy males who were listening to calm classical music and found their mean RMR to be 5708.07 with a standard deviation of 992.05.
At the $\alpha = 0.05$ level of significance, is there evidence to conclude that the mean RMR of males listening to calm classical music is different than 5710 kJ/day?

33. People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: $n = 40$ and $\bar{x} = 172.55$ *lb.*, and $\sigma = 26.33$ *lb*. Do not assume that the value of σ is known. Use these results to test the claim that men have a mean weight greater than 166.3 *lb.*, which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and the traditional method.

Section 3.8 – Hypothesis Tests for a Population Standard Deviation

Objective

Test a claim about a population standard deviation σ (or population variance σ^2) by using a formal method of hypothesis testing.

Notation

n = Sample size
s = sample standard deviation
s^2 = sample variance

σ = claimed value of the population standard deviation
σ^2 = claimed value of the population variance

Requirements for Testing Claims About a Population Mean (with σ Not Known)

1. The sample is a simple random sample.
2. The population has a normal distribution. (This is a much stricter requirement than the requirement of a normal distribution when testing claims about means.)

Chi-Square Distribution

If a simple random sample of size n is obtained from a normally distributed population with mean μ and standard deviation σ, then

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

where s^2 is a sample variance has a chi-square distribution with $n - 1$ degrees of freedom.

Properties of Chi-Square Distribution

1. It is ***not symmetric***.
2. The shape of the chi-square distribution depends on the degrees of freedom, just as with Student's t-distribution.
3. As the number of degrees of freedom increases, the chi-square distribution becomes more nearly symmetric.
4. The values of χ^2 are nonnegative (greater than or equal to 0).

Caution

The χ^2 test of this section is not *robust* against a departure from normality, meaning that the test does not work well if the population has a distribution that is far from normal. The condition of a normally distributed population is therefore a much stricter requirement in this section

Chi-Square Distribution Table

Chi-Square Distribution Table is based on cumulative areas from the right (unlike the entries in Standard Normal Distribution Table, which are cumulative areas from the left). Critical values are found in Chi-Square $\left(\chi^2\right)$ Distribution Table by first locating the row corresponding to the appropriate number of degrees of freedom (where df = n –1). Next, the significance level α is used to determine the correct column. The following examples are based on a significance level of $\alpha = 0.05$, but any other significance level can be used in a similar manner.

Right-tailed test: Because the area to the right of the critical value is 0.05, locate 0.05 at the top of Chi-Square Distribution Table.

Left-tailed test: With a left-tailed area of 0.05, the area to the right of the critical value is 0.95, so locate 0.95 at the top of Chi-Square Distribution Table.

Two-tailed test: Unlike the normal and Student *t* distributions, the critical values in this χ^2 test will be two different positive values (instead of something like ± 1.96). Divide a significance level of 0.05 between the left and right tails, so the areas to the right of the two critical values are 0.975 and 0.025, respectively. Locate 0.975 and 0.025 at the top of Chi-Square Distribution Table

Testing Hypotheses about a Population Variance or Standard Deviation

To test hypotheses about the population variance or standard deviation, we can use the following steps, provided that

- The sample is obtained using simple random sampling.
- The population is normally distributed.

***Step* 1**: Determine the null and alternative hypotheses. The hypotheses can be structured in one of three ways:

Two–Tailed	***Left–Tailed***	***Right–Tailed***
$H_0: \sigma = \sigma_0$	$H_0: \sigma = \sigma_0$	$H_0: \sigma = \sigma_0$
$H_1: \sigma \neq \sigma_0$	$H_1: \sigma < \sigma_0$	$H_1: \sigma > \sigma_0$

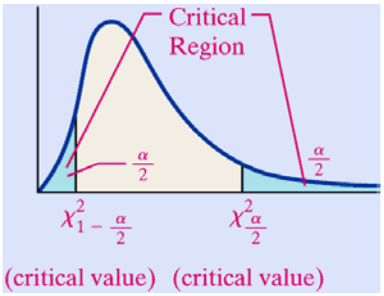
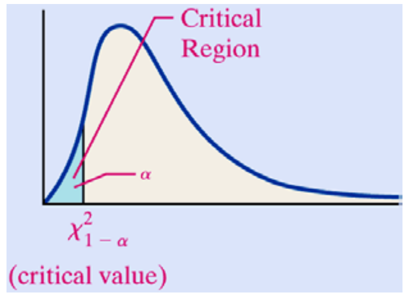
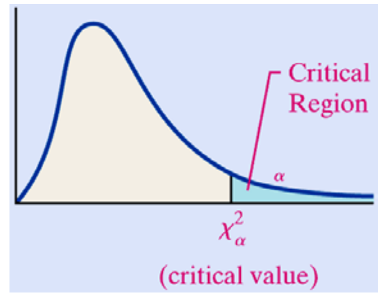
σ_0 is assumed value of the population standard deviation.

***Step* 2:** Select a level of significance, α, based on the seriousness of making a Type I error.

Step* 3**: Compute the ***test statistic $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

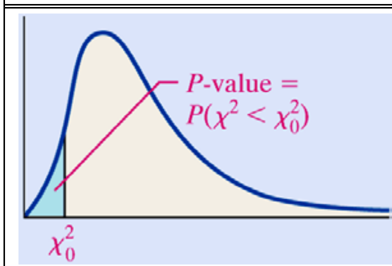
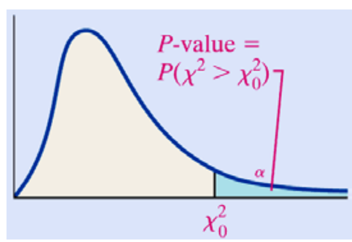
Classical Approach

Determine the critical value using n – 1 degrees of freedom.

Two–Tailed	*Left–Tailed*	*Right–Tailed*
$\chi_0^2 < \chi^2_{1-\alpha/2}$ or $\chi_0^2 > \chi^2_{\alpha/2}$ *Reject the null hypothesis*	$\chi_0^2 < \chi^2_{1-\alpha/2}$ *Reject the null hypothesis*	$\chi_0^2 > \chi^2_{\alpha/2}$ *Reject the null hypothesis*
		

P-Value Approach

To approximate the *P*-value for a left- or right-tailed test by determining the area under the chi-square distribution with $n - 1$ degrees of freedom to the left (for a left-tailed test) or right (for a right-tailed test) of the test statistic. For two-tailed tests, it is recommended that technology be used to find the *P*-value or obtain a confidence interval.

Left–Tailed	*Right–Tailed*
	

***Step* 4:** If the *P*-value < α, reject the null hypothesis.

***Step* 5:** State the conclusion.

CAUTION!

The procedures in this section are not robust.

Therefore, if analysis of the data indicates that the variable does not come from a population that is normally distributed, the procedures presented in this section are not valid.

Example

A can of 7-Up states that the contents of the can are 355 *ml*. A quality control engineer is worried that the filling machine is miscalibrated. In other words, she wants to make sure the machine is not under- or over-filling the cans. She randomly selects 9 cans of 7-Up and measures the contents. She obtains the following data.

351	360	358	356	359	358	355	361	352

We assumed the population standard deviation is 3.2. Test the claim that the population standard deviation, σ, is greater than 3.2 ml at the α = 0.05 level of significance.

Solution

Step 1: $H_0 : \sigma = 3.2$ *vs.* $H_1 : \sigma \neq 3.2$ This is a right-tailed test.

Step 2: The level of significance is α = 0.05.

Step 3: The sample standard deviation is computed $s = 3.464$. The test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(9-1)(3.464)^2}{3.2^2} \underline{= 9.374}$$

Classical Approach

Step 4: Since this is a right-tailed test, we determine the critical value at the α = 0.05 level of significance with 9 – 1 = 8 degrees of freedom to be $\chi^2_{0.05} = 15.507$.

Step 5: Since the test statistic, $\chi^2_0 = 9.374$, is less than the critical value, 15.507, we fail to reject the null hypothesis.

P-Value Approach

Step 4: Since this is a right-tailed test, the *P*-value is the area under the χ^2 distribution with $9 - 1 = 8$ to the right of the test statistic $\chi^2_0 = 9.374$. That is

$$P - value = P\left(\chi^2_0 > 9.37\right) > 0.1$$

Step 5: Since the *P*-value is greater than the level of significance, we fail to reject the null hypothesis.

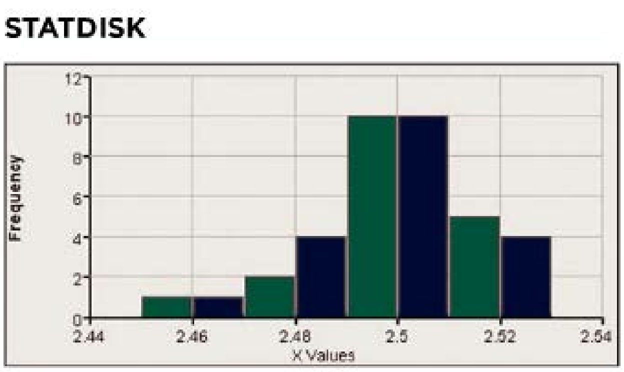
***Step* 6**: There is insufficient evidence at the α = 0.05 level of significance to conclude that the standard deviation of the can content of 7-Up is greater than 3.2 *ml*.

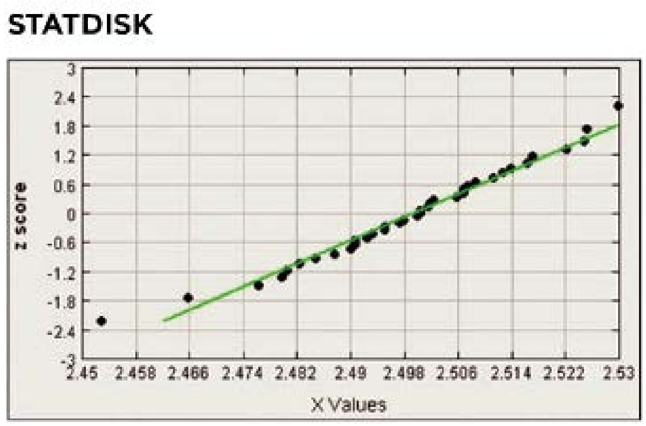
Example

A common goal in business and industry is to improve the quality of goods or services by reducing variation. Quality control engineers want to ensure that a product has an acceptable mean, but they also want to produce items of consistent quality so that there will be few defects. If weights of coins have a specified mean but too much variation, some will have weights that are too low or too high, so that vending machines will not work correctly (unlike the stellar performance that they now provide). Consider the simple random sample of the 37 weights of post-1983 pennies listed in Data Set 20 in Appendix B. Those 37 weights have a mean of 2.49910 g and a standard deviation of 0.01648 g. U.S. Mint specifications require that pennies be manufactured so that the mean weight is 2.500 g. A hypothesis test will verify that the sample appears to come from a population with a mean of 2.500 g as required, but use a 0.05 significance level to test the claim that the population of weights has a standard deviation less than the specification of 0.0230 g.

Solution

Requirements are satisfied: simple random sample; and STATDISK generated the histogram and quantile plot - sample appears to come from a population having a normal distribution.





Step 1: Express claim as $\sigma < 0.0230$ g

Step 2: If $\sigma < 0.0230$ g is false, then $\sigma \geq 0.0230$ g

Step 3: $\sigma < 0.0230$ g does not contain equality so it is the alternative hypothesis; null hypothesis is $\sigma = 0.0230$ g

H_0: $\sigma = 0.0230$ g

H_1: $\sigma < 0.0230$ g

Step 4: significance level is $\alpha = 0.05$

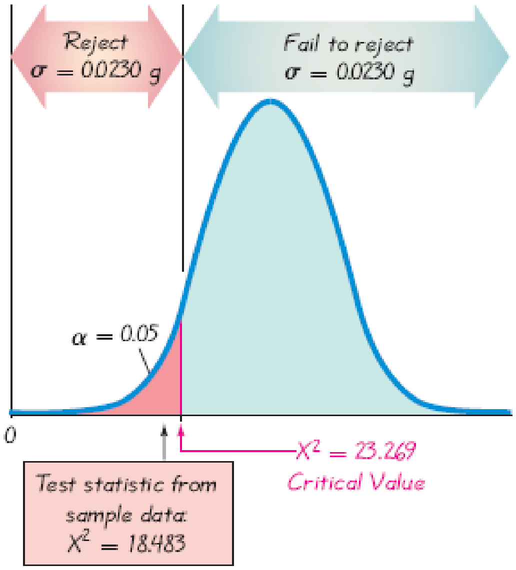
Step 5: Claim is about σ so use chi-square

Step 6: The test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(37-1)(0.01648)^2}{0.0230^2} = 18.483$$

The critical value from Chi-Square $\left(\chi^2\right)$ Distribution Table corresponds to 36 degrees of freedom and an "area to the right" of 0.95 (based on the significance level of 0.05 for a left-tailed test). Chi-Square $\left(\chi^2\right)$ Distribution Table does not include 36 degrees of freedom, but Chi-Square $\left(\chi^2\right)$ Distribution Table shows that the critical value is between 18.493 and 26.509. (Using technology, the critical value is 23.269.)

Step 7: Because the test statistic is in the critical region, reject the null hypothesis.



✓ There is sufficient evidence to support the claim that the standard deviation of weights is less than 0.0230 g. It appears that the variation is less than 0.0230 g as specified, so the manufacturing process is acceptable.

Exercises *Section* 3.8 – Hypothesis Tests for a Population Standard Deviation

1. There is a claim that the lengths of men's hands have a standard deviation less than 200 mm. You plan to test that claim with a 0.01 significance level by constructing a confidence interval. What level of confidence should be used for the confidence interval? Will the conclusion based on the confidence interval be the same as the conclusion based on a hypothesis test that uses the traditional method or the *P*-value method?

2. There is a claim that daily rainfall amounts in Boston have a standard deviation equal to 0.25 in. Sample data show that daily rainfall amounts are from a population with a distribution that is very far from normal. Can the use of a very large sample compensate for the lack of normality, so that the methods of this section can be used for the hypothesis test?

3. There is a claim that men have foot breaths with a variance equal to 36 mm^2. Is a hypothesis test of the claim that the variance is equal to 36 mm^2 equivalent to a test of the claim that the standard deviation is equal to 6 *mm*.

4. Given: $H_1: \sigma \neq 696\ g, \quad \alpha = 0.05, \quad n = 25, \quad s = 645\ g$, Find

a) Find the test statistic
b) Find critical value(s)
c) Find *P*-value limits
d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

5. Given: $H_1: \sigma < 29\ lb, \quad \alpha = 0.05, \quad n = 8, \quad s = 7.5\ lb$, Find

a) Find the test statistic
b) Find critical value(s)
c) Find *P*-value limits
d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

6. Given: $H_1: \sigma > 3.5\ \text{min}, \quad \alpha = 0.01, \quad n = 15, \quad s = 4.8\ \text{min}$, Find

a) Find the test statistic
b) Find critical value(s)
c) Find *P*-value limits
d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

7. Given: $H_1: \sigma \neq 0.25, \quad \alpha = 0.01, \quad n = 26, \quad s = 0.18$, Find

a) Find the test statistic
b) Find critical value(s)
c) Find *P*-value limits
d) Determine whether there is sufficient evidence to support the given alternative hypothesis.

8. A simple random sample of 40 men results in a standard deviation of 11.3 beats per minute. The normal range of pulse rates of adults is typically given as 60 to 100 beats per minute. If the range rule of thumb is applied to that normal range, the result is a standard deviation of 10 beats per minute. Use the sample results with a 0.05 significance level to test the claim that rates of men have a standard deviation greater than 10 beats per minute.

9. A simple random sample of 25 filtered 100 mm cigarettes is obtained, and the tar content of each cigarette is measured. The sample has a standard deviation of 3.7 mg. Use a 0.05 significance level to test the claim that the tar content of filtered 100 mm cigarettes has a standard deviation different from 3.2 mg, which is the standard deviation for unfiltered king size cigarettes.

10. When 40 people used the Weight Watchers diet for one year, their weight losses had a standard deviation of 4.9 lb. Use 0.01 significance level to test the claim that the amounts of weight loss have a standard deviation equal to 6.0 lb., which appears to be the standard deviation for the amounts of weight loss with the Zone diet.

11. Tests in the statistic classes have scores with a standard deviation equal to 14.1. One of the last classes has 27 test scores with a standard deviation of 9.3. Use a 0.01 significance level to test the claim that this class has less variation than other past classes. Does a lower standard deviation suggest that this last class is doing better?

12. A simple random sample of pulse rates of 40 women results in a standard deviation of 12.5 *beats/min*. The normal range of pulse rates of adults is typically given as 60 to 100 *beats/min*. If the range rule of thumb is applied to that normal range, the result is a standard deviation of 10 *beats/min*. Use the sample results with a 0.05 significance level to test the claim that pulse rates of women have a standard deviation equal to 10 *beats/min*.

13. Listed below are the playing times (in seconds) of songs that were popular at the time of this writing. Use a 0.05 significance level to test the claim that the songs are from a population with a standard deviation less than on minute.

448 242 231 246 246 293 280 227 244 213 262 239 213 258 255 257

14. Find the critical value or values of χ^2 based on the given information

a) $H_0 : \sigma = 8.0, \quad \alpha = 0.01, \quad n = 10$

b) $H_1 : \sigma > 3.5, \quad \alpha = 0.05, \quad n = 14$

c) $H_1 : \sigma < 0.14, \quad \alpha = 0.10, \quad n = 23$

d) $H_1 : \sigma \neq 9.3, \quad \alpha = 0.05, \quad n = 28$