Definition

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

Definition

The sample proportion $\hat{p}$ is the <u>best point estimate</u> of the population proportion p.

Definition

A confidence interval (or interval estimate) is a range (or an interval) of values used to estimate the true value of a population parameter.

A confidence interval is sometimes abbreviated as CI.

Example (continued)

In a recent poll, 70% of 1501 randomly selected adults said they believed in global warming.

The sample proportion $\hat{p}$ = 0.70 is the best estimate of the population proportion p.

A 95% confidence interval for the unknown population parameter is

$$0.677 < p < 0.723$$

What does it mean, exactly?

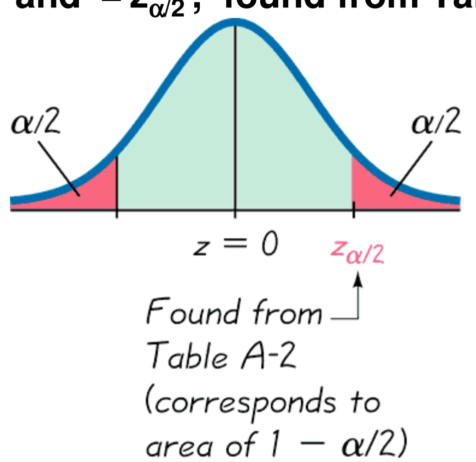
Interpreting a Confidence Interval

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, then 95% of them would actually contain the value of the population proportion p.

Critical Values

The z scores separate the middle interval (likely values) from the tails (unlikely values). They are $z_{\alpha/2}$ and $-z_{\alpha/2}$, found from Table A-2.



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.

Notation for Critical Value

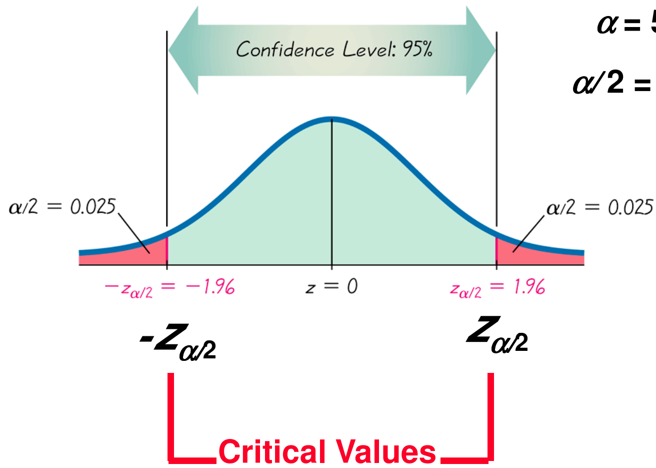
The critical value $z_{\alpha/2}$ separates an area of $\alpha/2$ in the right tail of the standard normal distribution. The value of $-z_{\alpha/2}$ separates an area of $\alpha/2$ in the left tail.

The subscript $\alpha/2$ is simply a reminder that the z score separates an area of $\alpha/2$ in the tail.

Finding $z_{\alpha/2}$ for a 95% Confidence Level

$\alpha = 5\%$

$\alpha/2 = 2.5\% = .025$



Definition

Margin of error, denoted by E, is the maximum likely difference (with probability $1 - \alpha$, 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.**

Confidence Interval for a Population Proportion p

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

where

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

Sample Size

Suppose we want to collect sample data in order to estimate some population proportion. The question is how many sample items must be obtained?

Sample Size for Estimating Proportion p

When an estimate of $\hat{p}$ is known:

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

When no estimate of $\hat{p}$ is known:

$$n = \frac{(z_{\alpha/2})^2\, 0.25}{E^2}$$

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.

Examples:

n=310.67 round up to 311
n=310.23 round up to 311
n=310.01 round up to 311

Example:

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.

Example:

a) Use $\hat{p} = 0.73$ and $\hat{q} = 1 - \hat{p} = 0.27$

$\alpha = 0.05$ so $z_{\alpha/2} = 1.96$

$E = 0.03$

$$n = \frac{(z_{\alpha/2})^2 \cdot \hat{p}\hat{q}}{E^2} = \frac{(1.96)^2 \cdot (0.73)(0.27)}{(0.03)^2} = 841.3104 = 842$$

To be 95% confident that our sample percentage is within three percentage points of the true percentage for all adults, we should obtain a random sample of 842 adults.

Example:

b) Use $\alpha = 0.05$ so $z_{\alpha/2} = 1.96$

$E = 0.03$

$$n = \frac{(z_{\alpha/2})^2 \cdot 0.25}{E^2} = \frac{(1.96)^2 \cdot 0.25}{(0.03)^2} = 1067.1111 = 1068$$

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

Section 7.3: Estimation of a population mean μ (σ is known)

In this section we cover methods for estimating a population mean. In addition to knowing the values of the sample data or statistics, we must also know the value of the population standard deviation, σ.

Point Estimate of the Population Mean

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

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

μ = population mean

σ = population standard deviation

$\bar{x}$ = 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 to check:

1. The value of the population standard deviation σ is known.

2. Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$. (Just like in the Central Limit Theorem.)

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

$$\bar{x} - E < \mu < \bar{x} + E \quad \text{where} \quad E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

or $\bar{x} \pm E$

or $(\bar{x} - E, \bar{x} + E)$

Definition

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

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 page 347

EXAMPLE 1 **Weights of Men** 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)?

Example page 347

Since :

1- Sample is random

2- the value of σ is known

3 – n >30

i.e. Conditions are satisfied

Then:

a) The sample mean 172.55 is the best point estimate for the population mean

Example page 347

b) 0.95 confidence level, $\alpha = 0.05$ then

$z_{\alpha/2} = 1.96$

$\sigma = 26$. and $n = 40$

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{26}{\sqrt{40}} = 8.0574835$$

$\bar{x} = 172.55$ and $E = 8.0574835$

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

$$164.49 < \mu < 180.61$$

Example page 347

C) Based on the confidence interval, 166.3 lb is a possible mean weight.

However the best point estimate 172.55 lb suggests that the mean weight of the men should be greater.

Since it is result in lives lost, then more analysis should be done

Finding a Sample Size for Estimating a Population Mean

μ = population mean

σ = population standard deviation

$\bar{x}$ = 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

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

Round-Off Rule for Sample Size n

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

Example:

Assume that we want to estimate the mean IQ score for the population of statistics students, with a standard deviation of 15. 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?

$\alpha = 0.05$
$\alpha/2 = 0.025$
$z_{\alpha/2} = 1.96$
$E = 3$
$\sigma = 15$

$$n = \left[\frac{1.96 \cdot 15}{3}\right]^2 = 96.04 = 97$$

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.

Section 7-5 Estimating a Population Variance

This section covers the estimation of a population variance σ^2 and standard deviation σ.

Estimator of σ^2

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

Estimator of σ

The **sample standard deviation** s is a commonly used **point estimate of** σ.

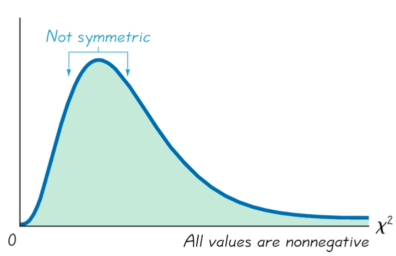
Construction of confidence intervals for σ^2

We use the **chi-square distribution**, denoted by Greek character χ^2 (pronounced *chi-square*).

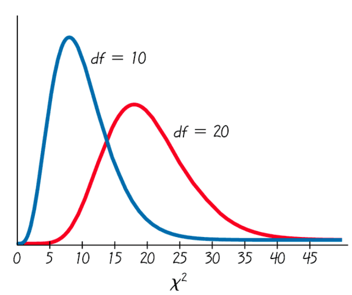
Properties of the Chi-Square Distribution

1. The chi-square distribution is **not symmetric**, unlike the normal and Student *t* distributions.

degrees of freedom = $n - 1$



Chi-Square Distribution



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

Properties of the Chi-Square Distribution

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

3. The chi-square distribution is different for each number of degrees of freedom, which is df = $n - 1$.

   In Table A-4, 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.

Example

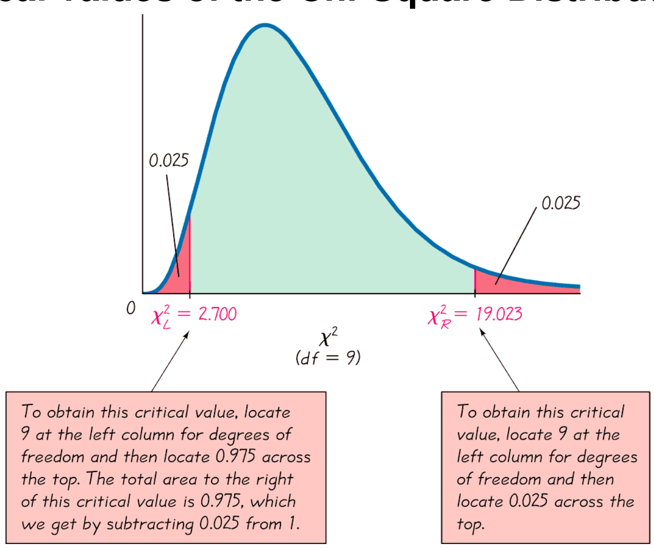
A sample of ten voltage levels is obtained.

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

Critical Values of the Chi-Square Distribution



Chi-Square Distribution

TABLE A-4 Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of the Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188

Confidence Interval for Estimating a Population Variance

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

Confidence Interval for Estimating a Population Standard Deviation

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

Confidence Interval for Estimating a Population Standard Deviation

For a sample of 10 voltages:

123.3 , 123.5, 123.7, 123.4, 1236, 123.5, 123.5, 123.4, 123.6, 123.8

The standard deviation s= 0.15

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

$$\frac{(10-1)(0.15)^2}{19.023} < \sigma^2 < \frac{(10-1)(0.15)^2}{2.700}$$

$$0.010645 < \sigma^2 < 0.075000.$$

$$0.10 \text{ volt} < \sigma < 0.27 \text{ volt}.$$