

Course Information

- [Schedule and Homework](#)
- [Course Requirements](#)
- [Instructor's Information](#)
- [Links](#)
- [Main Page](#)

Topics

1. [Introduction](#)
2. [Summarizing Data: listing and grouping](#)
3. [Summarizing Data: Statistical Descriptions](#)
4. [Possibilities and Probabilities](#)
5. [Some rules of probability](#)
6. [Probability Distributions](#)
7. [Normal Distributions](#)
8. [Sampling Distributions](#)
9. [Problems of Estimation](#)
10. [Hypothesis Tests](#)
11. [Tests based on count data](#)
12. [Regression](#)
13. [Non-parametric Tests](#)

Projects

1. [Stock return distribution](#)

- [Forum](#)

[edit](#) [sec](#) [src](#) [prt](#)

[Site Manager](#)

Chapter 9. Problems of Estimation

[Fold](#)

Table of Contents

- [Point estimates](#)
- [Point estimates](#)
- [Confidence Level](#)
- [Margin of Error](#)
- [Sample size](#)
- [Confidence intervals \(CI\)](#)
- [CI for the mean, for large samples \(when sd= \$\sigma\$ is known\)](#)
- [CI for the mean, for small samples \(when sd= \$\sigma\$ is not known\)](#)
- [T-distribution \(Student's T\)](#)
- [t-critical value](#)
- [Margin of error \(for the mean, small sample size\)](#)
- [Confidence Interval](#)
- [CI for a proportion \(large samples\).](#)
- [Point estimate](#)
- [Confidence Intervals Summary table](#)

Point estimates

When trying to estimate a population parameter with a sample statistic, it makes sense to use the single values of the sample mean \bar{x} , the sample standard deviation s or the sample variance s^2 to estimate the corresponding population parameters, the population mean μ , the population standard deviation σ or the population variance σ^2 , respectively.

Point estimates

We would then say that for example, the **point estimate for the population mean** $\hat{\mu} = \bar{x}$, where the \wedge on top of the population parameter μ indicates that that is the parameter we are estimating; i.e., the \wedge signifies that that value is an estimate (also called estimator). In words, the expression

$$\hat{\mu} = \bar{x} \tag{1}$$

indicates that we are using \bar{x} as the point estimate to estimate μ .

However, saying that for example the average age of the population of students is estimated using the sample mean, which is 22.4 years old, does not give enough information. We might be interested in knowing the size of the sample n , the standard deviation of the sample s . In fact, we know for a fact that the probability that the point estimate is exactly equal to the actual population parameter is virtually zero. For that reason, an interval (called confidence interval) might be of more use than a point estimate.

Confidence Level

To obtain such a confidence interval, the first thing to determine is a **confidence level** $1-\alpha$. This is a fixed probability that our estimate is within a certain range, which is chosen in advance. Oftentimes, the confidence level is 90%, 95%, 98%, or 99%. However, the higher the confidence level, the wider our range is, which could lead to some useless results. Therefore, a balance must be stroke between a high confidence level be stricken.

Margin of Error

When estimating the population mean with the sample mean and when the population standard deviation σ is known, the **maximum error of the estimate** for a given confidence level $1-\alpha$, also called the margin of error for the mean is

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} . \tag{2}$$

Sample size

Solving the equation above for n as the unknown allows to find the required sample size given a fixed desired margin of error.

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E} \right)^2 \tag{3}$$

Confidence intervals (CI)

CI for the mean, for large samples (when $sd=\sigma$ is known)

Once the margin of error has been determined, creating a confidence interval for the mean is easy. The lower bound of the interval is then

$\bar{x} - E$, while the upper bound of the interval is $\bar{x} + E$.

Therefore, the confidence interval for the mean with **1 - α** confidence level is

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \tag{4}$$

1 - α is also called the degree of confidence, and it represents the probability that for a large random sample this interval will contain the population mean μ .

CI for the mean, for small samples (when $sd=\sigma$ is not known)

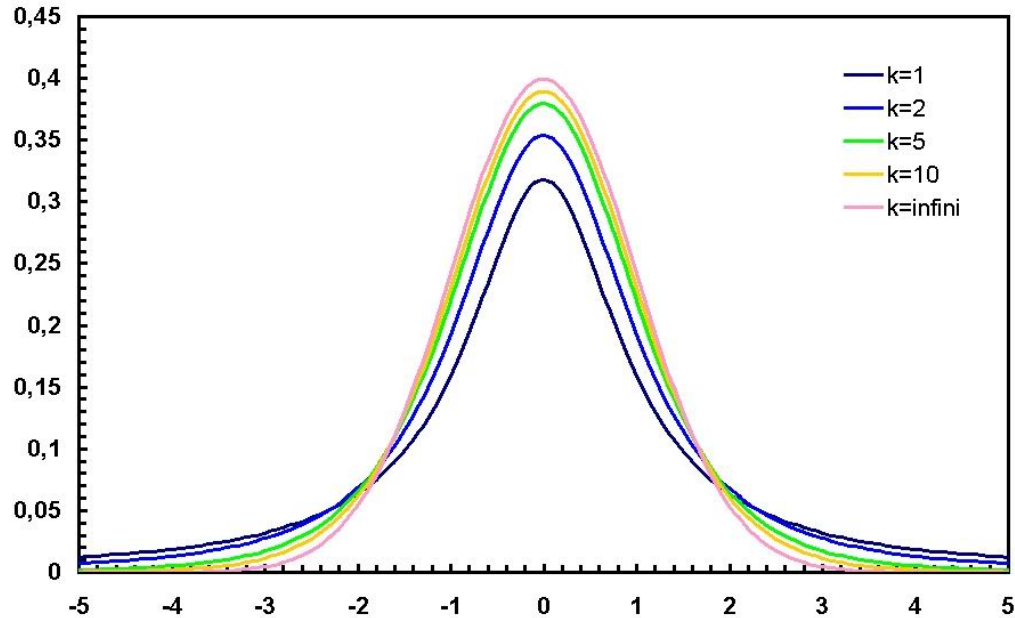
When we do not know σ , the population standard deviation, or when the sample is small and hence we cannot assume that $s = \sigma$, then we need to use another probability distribution called the **Student t-distribution**, instead of the normal distribution to find the critical value.

T-distribution (Student's T)

The t-distribution is

- centered at 0 (like Z)
- symmetrical (like Z)
- has "heavier" tails (more probability for large values and less probability around 0 than the standard normal distribution)
- depends on a parameter called **degrees of freedom df**
- the larger the degrees of freedom, the closer it is to the standard normal distribution Z (and in the limit as $n \rightarrow \infty$ it actually becomes Z).

Some example graphs are given below:



t-critical value

To find the t-critical value, we will need to read from a t-table (one is available [here](#)). Once you know the confidence level, $1-\alpha$ (for example 90% confidence), assign probability $\alpha/2$ (5% in our example) to the right tail (this determines which column to read from). Then, we also need to know the degrees of freedom, which determine the row to read from. In the case of a

confidence interval for the mean for a small sample is $df=n-1$.

For example, assume that our sample had sample size $n=10$.

Then, reading from the table under $t_{0.050}$ which represents the right tail probability of 5%, and on the row with $df=n-1=10-1=9$, we obtain

df	...	$t_{0.050}$
...
9	...	1.833

Therefore, the critical value $t_{\alpha/2}=t_{0.050}=1.833$.

Margin of error (for the mean, small sample size)

The margin of error in this case is

$$E = t_{\alpha/2} \frac{s}{\sqrt{n}} . \tag{5}$$

Confidence Interval

Once the margin of error has been determined, creating a confidence interval for the mean is easy. The lower bound of the interval is then

$\bar{x} - E$, while the upper bound of the interval is $\bar{x} + E$.

Therefore, the confidence interval for the mean with $1 - \alpha$ confidence level is

$$\bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n}} \tag{6}$$

CI for a proportion (large samples).

In the case of an estimate for a proportion, the

Point estimate

for the population proportion is $\hat{p}=x/n$,
where x is the number
Once the margin of error has been determined, creating a confidence interval for the mean is easy. The lower bound of the interval is then $\bar{x} - E$, while the upper bound of the interval is $\bar{x} + E$.
Therefore, the confidence interval for the mean with $1 - \alpha$ confidence level is

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

(7)

$1 - \alpha$ is also called the degree of confidence, and it represents the probability that for a large random sample this interval will contain the population mean μ .

Confidence Intervals Summary table

CI for	details	distribution, df	Margin of Error	Confidence interval
Mean μ	$n \geq 30, \sigma$ known	z (st. normal dist)	$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$	$\bar{x} - E < \mu < \bar{x} + E$
Mean μ	$n < 30, \sigma$ not known	t (df = n-1)	$E = t_{\alpha/2} \frac{s}{\sqrt{n}}$	$\bar{x} - E < \mu < \bar{x} + E$
Proportion p	$n > 30$	z	$E = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$	$\frac{\hat{x}}{n} - E < p < \frac{\hat{x}}{n} + E$
Standard deviation σ	large sample	z	n.a.	$\frac{s}{1 + \frac{z_{\alpha/2}}{\sqrt{2n}}} < \sigma < \frac{s}{1 - \frac{z_{\alpha/2}}{\sqrt{2n}}}$
Standard deviation σ	small sample	χ^2 , df = n-1	n.a.	$\sqrt{\frac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_L}}$

page revision: 22, last edited: 29 Apr 2009, 15:51 (3547 days ago)

[Edit](#) [Tags](#) [History](#) [Files](#) [Print](#) [Site tools](#) [+ Options](#)

Powered by [Wikidot.com](#)

[Help](#) | [Terms of Service](#) | [Privacy](#) | [Report a bug](#) | [Flag as objectionable](#)

Unless otherwise stated, the content of this page is licensed under [Creative Commons Attribution-ShareAlike 3.0 License](#)