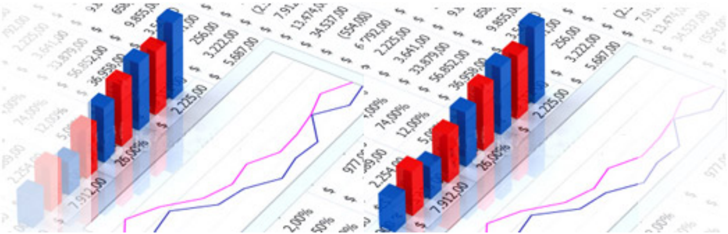
MTMS.01.099 Mathematical Statistics

Lecture 8. Confidence intervals

Tõnu Kollo



Fall 2017

Interval Estimation

- A point estimate, because it is a single number, does not give sufficient information about an unknown parameter – provides no information about the precision and reliability of the estimate
- It is often more appropriate to use an *interval estimate*, or, as we say, a *confidence interval*.

Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ be a sample from a distribution $X \sim P_X$ that depends on the parameter θ and $\boldsymbol{X} = (X_1, \ldots, X_n)$ – corresponding theoretical sample.

Definition

An interval I_θ that covers θ with probability $1 - \alpha$ is called a **confidence interval** for θ with *confidence level* $1 - \alpha$.

Interval Estimation (2)

Given definition means that there are functions of theoretical sample $a_1(\boldsymbol{X})$ and $a_2(\boldsymbol{X})$ so that

$$P(a_1(\boldsymbol{X}) \leq \theta \leq a_2(\boldsymbol{X})) = 1 - \alpha.$$

Then values of these random variables $a_1(\boldsymbol{x})$ and $a_2(\boldsymbol{x})$ are called **lower and upper confidence limits** and

$$I_\theta = (a_1(\boldsymbol{x}), a_2(\boldsymbol{x}))$$

is a confidence interval at the confidence level $1 - \alpha$.

- The main problem is to select the two functions.
- We seek functions that make the interval as short as possible to make the interval estimate precise.

Interval Estimation (3)

Requirements for confidence interval:

- as narrow as possible
- confidence level as high as possible (closer to 1)

A very wide confidence interval gives the message that there is a great deal of uncertainty concerning the value of what we are estimating.

If confidence limits $a_1(\boldsymbol{x})$ and $a_2(\boldsymbol{x})$ are finite, the interval is then called **two-sided**, otherwise the interval is called **one-sided** $(-\infty, a_2), (a_1, \infty)$.

In principle, the probability $1 - \alpha$ can be fixed at any value. The classical confidence levels are

$$1 - \alpha = \begin{cases} 0.90 \\ 0.95 \\ 0.99 \end{cases}$$

Interval Estimation. Example 1

Recall from Chapter 8.4 from the book the following formulas: if $X \sim N(\theta, \sigma^2)$, then we are interested in such constants λ_α and $\lambda_{\alpha/2}$ that

$$P(X < \theta + \lambda_\alpha \sigma) = 1 - \alpha$$

$$P(\theta - \lambda_{\alpha/2}\sigma < X < \theta + \lambda_{\alpha/2}\sigma) = 1 - \alpha.$$

What these constants are?

$$P(X < \theta + \lambda_\alpha \sigma) = P(X - \theta < \lambda_\alpha \sigma) = P(\frac{X - \theta}{\sigma} < \lambda_\alpha) = \Phi(\lambda_\alpha) = 1 - \alpha.$$

λ_α is the $(1 - \alpha)$-quantile of $N(0, 1)$ and is called also the complement α-quantile

Interval Estimation. Example 1

Example

A research worker measures a quantity whose unknown value is θ and gets the value x. Hence x is an observation on X, where $X \sim N(\theta, \sigma^2)$. We want to construct a two-sided confidence interval for θ at the confidence level 0.95 (more briefly – a 95% confidence interval).

As $X \sim N(\theta, \sigma^2)$,

$$P(\theta - \lambda_{\alpha/2}\sigma < X < \theta + \lambda_{\alpha/2}\sigma) = 1 - \alpha,$$

where $\lambda_{\alpha/2}$ is $N(0,1)$ complement $\alpha/2$ quantile.

What's the value of $\lambda_{\alpha/2}$ at confidence level 95%?

Interval Estimation. Example 1 (3)

$X \sim N(\theta, \sigma^2)$, the probability is 0.95 that

$$\theta - 1.96\sigma < X < \theta + 1.96\sigma,$$

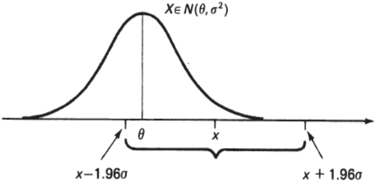
or, equivalently, that

$$X - 1.96\sigma < \theta < X + 1.96\sigma.$$

Hence the interval

$$I_\theta = (x - 1.96\sigma, x + 1.96\sigma)$$

is a 95% confidence interval.



Remarks

- Be careful when reading an equation such as $0.95 = P(X - 1.96\sigma < \theta < X + 1.96\sigma)$. This does not mean that θ is random with a 0.95 probability of falling between two values. The parameter θ is an unknown *constant*. Instead, it is the interval that is random, with a 95% probability of including θ.
- The statement "we are 95% confident" means that if we repeated the same process of drawing samples and computing intervals many times, then in the long run, 95% of the intervals would include θ.

Example 2

Example

An engineer tests the gas consumption of a random sample of 30 of his company cars ready to be sold. The 95% confidence interval for the average consumption of cars of all the cars is (7.8; 8.8) litres per 100 km. Evaluate the following statements:

1. We are 95% confident that the gas consumption per 100 km for cars in this company is between 7.8 and 8.8 litres.

2. 95% of all samples will give an average consumption between 7.8 and 8.8 litres.

3. There is a 95% chance that the true mean is between 7.8 and 8.8 litres.

Example continues

Solution:
1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.

Example continues

Solution:
1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.
This is not correct: a confidence interval is for a population parameter, and in this case the mean, not for individuals.

Example continues

Solution:
1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.
This is not correct: a confidence interval is for a population parameter, and in this case the mean, not for individuals.
2. 95% of all samples will give an average consumption between 7.8 and 8.8 litres/100 km.

Example continues

Solution:
1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.
This is not correct: a confidence interval is for a population parameter, and in this case the mean, not for individuals.
2. 95% of all samples will give an average consumption between 7.8 and 8.8 litres/100 km.
This is not correct: each sample will give rise to a different confidence interval and 95% of these intervals will contain the true mean.

Example continues

Solution:

1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.

This is not correct: a confidence interval is for a population parameter, and in this case the mean, not for individuals.

2. 95% of all samples will give an average consumption between 7.8 and 8.8 litres/100 km.

This is not correct: each sample will give rise to a different confidence interval and 95% of these intervals will contain the true mean.

3. There is a 95% chance that the true mean is between 7.8 and 8.8 litres/100 km.

Example continues

Solution:

1. We are 95% confident that the gas consumption for cars in this company is between 7.8 and 8.8 litres/100 km.

This is not correct: a confidence interval is for a population parameter, and in this case the mean, not for individuals.

2. 95% of all samples will give an average consumption between 7.8 and 8.8 litres/100 km.

This is not correct: each sample will give rise to a different confidence interval and 95% of these intervals will contain the true mean.

3. There is a 95% chance that the true mean is between 7.8 and 8.8 litres/100 km.

This is not correct: θ is not random. The probability that it is between 7.8 and 8.8 is 0 or 1.

General Method

- find a point estimate for θ, $\hat{\theta}(\boldsymbol{x})$ – function of a sample;
- find the density function of a corresponding point estimator $\hat{\boldsymbol{\theta}} = \hat{\theta}(\boldsymbol{X})$ (depends on θ);
- using the determined density function, find two quantities $b_1(\theta)$ and $b_2(\theta)$ such that

$$P(b_1(\theta) \leq \hat{\theta}(\boldsymbol{X}) \leq b_2(\theta)) = 1 - \alpha, \quad (*)$$

where α is fixed and known.

General Method (1)

- if $b_1(\theta)$ and $b_2(\theta)$ are strictly monotone functions, then the inverse functions $b_1^{-1}(\cdot)$ and $b_2^{-1}(\cdot)$ exist, and the inequality in $(*)$ can be rewritten

$$P(a_1(\hat{\theta}) \le \theta \le a_2(\hat{\theta})) = 1 - \alpha,$$

For example, if $b_1(\cdot)$ and $b_2(\cdot)$ are strictly increasing functions, then

$$b_1(\theta) \le \hat{\theta} \Rightarrow \theta \le b_1^{-1}(\hat{\theta}) \text{ and } \hat{\theta} \le b_2(\theta) \Rightarrow b_2^{-1}(\hat{\theta}) \le \theta,$$

so the lower confidence limit is $a_1(\hat{\theta}) = b_2^{-1}(\hat{\theta})$ and upper is $a_2(\hat{\theta}) = b_1^{-1}(\hat{\theta})$.

General Method (2)

- Since $\hat{\boldsymbol{\theta}} = \hat{\theta}(\boldsymbol{X})$, the confidence limits are also functions of theoretical sample,

$$P(a_1(\boldsymbol{X}) \leq \theta \leq a_2(\boldsymbol{X})) = 1 - \alpha.$$

- Substituting now theoretical sample with realised sample, $(\boldsymbol{X} = \boldsymbol{x})$, we get that $I_\theta = (a_1(\boldsymbol{x}), a_2(\boldsymbol{x}))$ is a CI for a parameter θ at the confidence level $1 - \alpha$.

General Method. Remark

Remark

Sometimes if the distribution of $\hat{\boldsymbol{\theta}}$ is unknown, but we know the distribution of its one-to-one function $g(\hat{\boldsymbol{\theta}})$, then we can find $b_1(\theta)$ and $b_2(\theta)$ such that

$$P(b_1(\theta) \leq g(\hat{\boldsymbol{\theta}}) \leq b_2(\theta)) = 1 - \alpha.$$

Usually it is possible to derive a confidence interval for θ from here.

General Method. Example

Example

Suppose we want to construct a two-sided CI (confidence interval) for average weight of Estonian men (aged 35-54) with confidence level 0.95. It is known that the standard deviation of weight (of Estonian men with given age) is 15kg. We have one randomly obtained observation, $x = 85$. Let us assume that weight has a theoretical distribution $N(\theta, 15^2)$, where θ is our unknown average. Construct the two-sided CI for average weight at the confidence level 0.95.

General Method. Example (1)

- find a point estimate for θ, $\hat{\theta}(\boldsymbol{x})$ – function of a sample:
  θ - average weight of men. Point estimate of θ is a sample mean, $\hat{\theta}(\boldsymbol{x}) = \overline{x} = x_1 = 85$.
- find the density function of a corresponding point estimator $\hat{\boldsymbol{\theta}} = \hat{\theta}(\boldsymbol{X})$ (depends on θ):
  Corresponding point estimator is X_1 and it is known that $X_1 \sim N(\theta, 15^2)$.
- using the determined density function, find two quantities $b_1(\theta)$ and $b_2(\theta)$ such that $P(b_1(\theta) \leq \hat{\theta}(\boldsymbol{X}) \leq b_2(\theta)) = 1 - \alpha$.

  $$P(\theta - \lambda_{\alpha/2}\sigma \leq X_1 \leq \theta + \lambda_{\alpha/2}\sigma) = 1 - \alpha \ \Rightarrow$$

  $b_1(\theta) = \theta - \lambda_{\alpha/2}\sigma$ and $b_2(\theta) = \theta + \lambda_{\alpha/2}\sigma$.

General Method. Example (2)

- if $b_1(\theta)$ and $b_2(\theta)$ are strictly monotone functions, then the inverse functions $b_1^{-1}(\cdot)$ and $b_2^{-1}(\cdot)$ exist, and the inequality in $(*)$ can be rewritten

$$P(a_1(\hat{\boldsymbol{\theta}}) \leq \theta \leq a_2(\hat{\boldsymbol{\theta}})) = 1 - \alpha,$$

For example, if $b_1(\cdot)$ and $b_2(\cdot)$ are strictly increasing functions, then

$$b_1(\theta) \leq \hat{\boldsymbol{\theta}} \Rightarrow \theta \leq b_1^{-1}(\hat{\boldsymbol{\theta}}) \text{ and } \hat{\boldsymbol{\theta}} \leq b_2(\theta) \Rightarrow b_2^{-1}(\hat{\boldsymbol{\theta}}) \leq \theta,$$

so the lower confidence limit is $a_1(\hat{\boldsymbol{\theta}}) = b_2^{-1}(\hat{\boldsymbol{\theta}})$ and upper is $a_2(\hat{\boldsymbol{\theta}}) = b_1^{-1}(\hat{\boldsymbol{\theta}})$.

How to find inverse function:	
1. Function y=f(x) is given	1. $b_1(\theta) = \theta - \lambda_{\alpha/2}\sigma$
2. Derive a variable x from it	2. $\theta = b_1(\theta) + \lambda_{\alpha/2}\sigma$
3. Switch x and y	3. $b_1^{-1}(\theta) = \theta + \lambda_{\alpha/2}\sigma$
	$\hat{\theta} = X_1$, $b_1^{-1}(\hat{\theta}) = X_1 + \lambda_{\alpha/2}\sigma$.

General Method. Example (3)

- Since $\hat{\boldsymbol{\theta}} = \hat{\theta}(\boldsymbol{X})$, the confidence limits are also functions of theoretical sample,

$$P(a_1(\boldsymbol{X}) \leq \theta \leq a_2(\boldsymbol{X})) = 1 - \alpha.$$

Deriving analogically $a_1(\hat{\theta}) = b_2^{-1}(\hat{\theta})$, we get

$$P(X_1 - \lambda_{\alpha/2}\sigma \leq \theta \leq X_1 + \lambda_{\alpha/2}\sigma) = 1 - \alpha$$

- Substituting now theoretical sample with implemented sample (realisations), $(X = x)$, we get that $I_\theta = (a_1(\boldsymbol{x}), a_2(\boldsymbol{x}))$ is a CI for a parameter θ with confidence level $1 - \alpha$.

$$I_\theta = (x_1 - \lambda_{\alpha/2}\sigma; x_1 + \lambda_{\alpha/2}\sigma) = (55.6; 114.4)$$