

Lecture 29

The confidence interval formulas for the mean in a normal distribution when σ is unknown

1 Introduction

In this lecture we will derive the formulas for the symmetric two-sided confidence interval and the lower-tailed confidence intervals for the mean in a normal distribution *when the variance σ^2 is unknown*. At the end of the lecture I assign the problem of proving the formula for the upper-tailed confidence interval. We will need the following theorem from probability theory. Recall that \bar{X} is the sample mean (the point estimator for the population mean μ) and S^2 is the sample variance, the point estimator for the unknown population variance σ^2 .

We will need the following theorem from Probability Theory.

Theorem 1. $(\bar{X} - \mu) / \frac{S}{\sqrt{n}}$ has *t-distribution with $n - 1$ degrees of freedom*.

2 The two-sided confidence interval formula

Now we can prove the theorem from statistics giving the required confidence interval for μ . Note that it is symmetric around \bar{X} . There are also asymmetric two-sided confidence intervals. We will discuss them later. This is one of the basic theorems that you have to learn how to prove.

Theorem 2. *The random interval $T = (\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}})$ is a $100(1 - \alpha)\%$ -confidence interval for μ .*

Proof. We are required to prove

$$P(\mu \in (\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}})) = 1 - \alpha.$$

We have

$$\begin{aligned}
LHS &= P\left(\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} < \mu, \mu < \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}\right) \\
&= P\left(\bar{X} - \mu < t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, -t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} < \bar{X} - \mu\right) \\
&= P\left(\bar{X} - \mu < t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{X} - \mu > -t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}\right) \\
&= P\left(\frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}} < t_{\alpha/2, n-1}, \frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}} > -t_{\alpha/2, n-1}\right) \\
&= P(T < t_{\alpha/2, n-1}, T > -t_{\alpha/2, n-1}) = P(-t_{\alpha/2, n-1} < T < t_{\alpha/2, n-1}) = 1 - \alpha
\end{aligned}$$

To prove the last equality draw a picture. \square

Once we have an actual sample x_1, x_2, \dots, x_n we obtain the observed value \bar{x} for the random variable \bar{X} and the observed value s for the random variable S . We obtain the observed value (an ordinary interval) $(\bar{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}})$ for the confidence (random) interval $(\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}})$. The observed value of the confidence (random) interval is also called the two-sided $100(1 - \alpha)\%$ confidence interval for μ .

3 The lower-tailed confidence interval

In this section we will give the formula for the lower-tailed confidence interval for μ .

Theorem 3. *The random interval $(-\infty, \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}})$ is a $100(1 - \alpha)\%$ -confidence interval for μ .*

Proof. We are required to prove

$$P(\mu \in (-\infty, \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}})) = 1 - \alpha.$$

We have

$$\begin{aligned}
LHS &= P(\mu < \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}}) = P(-t_{\alpha, n-1} \frac{S}{\sqrt{n}} < \bar{X} - \mu) \\
&= P(-t_{\alpha, n-1} < \frac{(\bar{X} - \mu)}{\frac{S}{\sqrt{n}}}) \\
&= P(-t_{\alpha, n-1} < T) \\
&= 1 - \alpha
\end{aligned}$$

To prove the last equality draw a picture - I want *you* to draw the picture on tests and the homework. \square

Once we have an actual sample x_1, x_2, \dots, x_n we obtain the observed value \bar{x} for the random variable \bar{X} the observed value s for the random variable S and the observed value $(-\infty, \bar{x} + t_{\alpha, n-1} \frac{s}{\sqrt{n}})$ for the confidence (random) interval $(-\infty, \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}})$. The observed value of the confidence (random) interval is also called the lower-tailed $100(1 - \alpha)\%$ confidence interval for μ .

The random variable $\bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}}$ or its observed value the number $\bar{x} + t_{\alpha, n-1} \frac{s}{\sqrt{n}}$ is often called a confidence *upper bound* for μ because

$$P(\mu < \bar{X} + t_{\alpha, n-1} \frac{S}{\sqrt{n}}) = 1 - \alpha.$$

4 The upper-tailed confidence interval for μ

Problem Prove the following theorem.

Theorem 4. *The random interval $(\bar{X} - t_{\alpha, n-1} \frac{S}{\sqrt{n}}, \infty)$ is a $100(1 - \alpha)\%$ confidence interval for μ .*

The random variable $\bar{X} - t_{\alpha, n-1} \frac{S}{\sqrt{n}}$ or its observed value the number $\bar{x} - t_{\alpha, n-1} \frac{s}{\sqrt{n}}$ is often called a confidence *lower bound* for μ because

$$P(\mu > \bar{X} - t_{\alpha, n-1} \frac{S}{\sqrt{n}}) = 1 - \alpha.$$