

Lecture 31: The prediction interval formulas for the next observation from a normal distribution when σ is known

1. Introduction

In this lecture we will derive the formulas for the symmetric two-sided prediction interval for the $n + 1$ --st observation and the upper-tailed prediction interval for the $n + 1$ -st observation from a normal distribution *when the variance σ^2 is known*. We will need the following theorem from probability theory that gives the distribution of the statistic $\bar{X} - X_{n+1}$.

Suppose that $X_1, X_2, \dots, X_n, X_{n+1}$ is a random sample from a normal distribution with mean μ and variance σ^2 . We assume μ is unknown but σ^2 is known.

Theorem 1

The random variable $\bar{X} - X_{n+1}$ has normal distribution with mean zero and variance $\frac{n+1}{n}\sigma^2$. Hence we find that the random variable

$Z = (\bar{X} - X_{n+1}) / \left(\sqrt{\frac{n+1}{n}}\sigma \right)$ has standard normal distribution.

2. The two-sided prediction interval formula

Now we can prove the theorem from statistics giving the required prediction interval for the next observation x_{n+1} in terms of n observations x_1, x_2, \dots, x_n . Note that it is symmetric around \bar{X} . This is one of the basic theorems that you have to learn how to prove. There are also asymmetric two-sided prediction intervals.

Theorem 2

The random interval $\left(\bar{X} - z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma, \bar{X} + z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma\right)$ is a $100(1 - \alpha)\%$ -prediction interval for x_{n+1} .

Proof.

We are required to prove

$$P\left(X_{n+1} \in \left(\bar{X} - z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma, \bar{X} + z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma\right)\right) = 1 - \alpha.$$

We have

$$\begin{aligned} \text{LHS} &= P\left(\bar{X} - z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma < X_{n+1}, X_{n+1} < \bar{X} + z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma\right) \\ &= P\left(\bar{X} - X_{n+1} < z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma\right) \\ &= P\left(\bar{X} - X_{n+1} < z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma, \bar{X} - X_{n+1} > -z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma\right) \\ &= P(Z < z_{\alpha/2}, Z > -z_{\alpha/2}) = P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha \end{aligned}$$

To prove the last equality draw a picture. □

Once we have an actual sample x_1, x_2, \dots, x_n we obtain the observed value $\left(\bar{x} - z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma, \bar{x} + z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma \right)$ for the prediction (random) interval $\left(\bar{X} - z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma, \bar{X} + z_{\alpha/2} \sqrt{\frac{n+1}{n}} \sigma \right)$. The observed value of the prediction (random) interval is also called the two-sided $100(1 - \alpha)\%$ prediction interval for x_{n+1} .

3. The upper-tailed prediction interval

In this section we will give the formula for the upper-tailed prediction interval for the next observation x_{n+1} .

Theorem 3

The random interval $(\bar{X} - z_{\alpha} \sqrt{n+1} n \sigma, \infty)$ is a $100(1 - \alpha)\%$ -prediction interval for the next observation x_{n+1} .

Proof

We are required to prove

$$P(X_{n+1} \in (\bar{X} - z_{\alpha} \sqrt{\frac{n+1}{n}} \sigma, \infty)) = 1 - \alpha.$$

Proof (Cont.)

We have

$$\text{LHS} = P\left(\bar{X} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma < X_{n+1}\right)$$

To prove the last equality draw a picture - I want you to draw the picture on tests and the final. □

Once we have an actual sample x_1, x_2, \dots, x_n we obtain the observed value

$\left(\bar{x} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma, \infty\right)$ of the upper-tailed prediction (random) interval

$\left(\bar{X} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma, \infty\right)$ The observed value of the upper-tailed prediction (random) interval is also called the upper-tailed $100(1 - \alpha)\%$ prediction interval for x_{n+1} .

The number random variable $\bar{X} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma$ or its observed value

$\bar{x} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma$ is often called a prediction *lower bound* for x_{n+1} because

$$P\left(\bar{X} - z_\alpha \sqrt{\frac{n+1}{n}}\sigma < X_{n+1}\right) = 1 - \alpha.$$