Lecture 24

niceguy

March 13, 2023

1 Confidence Intervals

Recall with n IID samples, an observec mean $\overline{X}$, a known variance σ^2 and statistic

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

By the central limit theorem, Z has $n(z; 0, 1)$. We set

$$z_\beta = -\Phi^{-1}(\beta)$$

$$\begin{aligned}
1 - \alpha &= P(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}) \\
&= P\left(\overline{X} - \frac{z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right) \\
&= P(\overline{X}_L \leq Z \leq \overline{X}_U)
\end{aligned}$$

1.1 Realised Confidence Interval

The real confidence interval is then

$$\left[\overline{x} - \frac{z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}, \overline{x} + \frac{z_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right]$$

1.2 One Sided Confidence Interval

There is also a one sided confidence interval

$$1 - \alpha = P(Z \leq z_\alpha)$$

where similarly, we set

$$z_\alpha = -\Phi^{-1}(\alpha)$$

and we have

$$1 - \alpha = P\left(\mu \leq \overline{X} + \frac{z_\alpha \sigma}{\sqrt{n}}\right)$$

2 Confidence Interval with Unknown Variance

If variance is not known, with normal samples, then setting

$$T = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}$$

where

$$S^2 = \frac{1}{n-1} \sum_i \left(X_i - \overline{X}\right)^2$$

Then T has a t distribution. For $\beta < 0.5$, letting $H(t)$ be the cumulative distirbution function, we can define t_β such that

$$t_\beta = H^{-1}(\beta)$$

Then similarly,

$$\begin{aligned}
1 - \alpha &= P\left(-t_{\frac{\alpha}{2}} \leq T \leq t_{\frac{\alpha}{2}}\right) \\
&= P\left(-t_{\frac{\alpha}{2}} \leq \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}}\right) \\
&= P\left(\overline{X} - \frac{t_{\frac{\alpha}{2}} S}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{t_{\frac{\alpha}{2}} S}{\sqrt{n}}\right)
\end{aligned}$$

Then given $\overline{x}$, its realisation is

$$\left[\overline{x} - \frac{t_{\frac{\alpha}{2}} S}{\sqrt{n}}, \overline{x} + \frac{t_{\frac{\alpha}{2}} S}{\sqrt{n}}\right]$$

Example 2.1. Let $n = 7$, with normal samples. The observed mean $\overline{x}$ is -3, and the observed variance $s^2 = 2$. Setting $\alpha = 0.1$ for a 90% confidence interval,

$$t_{0.05} = -1.9$$

so the realised confidence interval is

$$\left[-3 - 1.9 \times \frac{2}{\sqrt{7}}, -3 + 1.9 \times \frac{2}{\sqrt{7}}\right] = [-3.52, -2.48]$$

3 Standard Error

For samples $X_1, \ldots, X_n$, and

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

with normal distribution, $\overline{X}$ has a standard error of $\frac{\sigma}{\sqrt{n}}$. The width of the confidence interval is proportional to this.

4 Prediction Intervals

With samples $X_1, \ldots, X_n$ that are normal, let X_0 be a new observation. Now $\overline{X}$ is a good point estimation of X_0. The error is $X_0 - \overline{X}$, and the variance of the error is $\sigma^2 + \frac{\sigma^2}{n}$. For the statistic

$$Z = \frac{X_0 - \overline{X}}{\sigma\sqrt{1 + \frac{1}{n}}}$$

Z has a normal distribution, then

$$\begin{aligned} 1 - \alpha &= P(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}) \\ &= P\left(\overline{X} - z_{\frac{\alpha}{2}}\sigma\sqrt{1 + \frac{1}{n}} \leq X_0 \leq \overline{X} + z_{\frac{\alpha}{2}}\sigma\sqrt{1 + \frac{1}{n}}\right) \end{aligned}$$

Where

$$z_{\frac{\alpha}{2}} = -\Phi^{-1}\left(\frac{\alpha}{2}\right)$$