Lecture 37 (Review Lecture)

niceguy

April 13, 2023

1 Central Limit Theorem

We have a sample with data $x_1, \ldots, x_n$, which are **actual numbers**. There are also random variables $X_1, \ldots, X_n$, whose realisations are the data. They are independent and identically distributed. So they have the same (unknown) distribution.

2 Mean

The numeric value is

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

The random variable is

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

The true mean is μ, and the true standard deviation is σ. Then let

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

The Central Limit Theorem states that as $n \to \infty$, the distribution of Z approaches the normal distribution.

3 Confidence Interval

Letting Z be a normal distribution, we find a range where there is a probability of $x\%$ that the true mean lies in said region. Then defining

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

we can rearrange the terms to get

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

The confidence interval is then

$$\left(\overline{X} - z_{\frac{\alpha}{2}}, \overline{X} + z_{\frac{\alpha}{2}}\right)$$

4 T-Distribution

Given a **normal** population, with an unknown σ, which is estimated from the sample, then it is the **exact distribution**. The same procedure (using a t instead of a z) follows. We estimate the standard deviation by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Now if $n < 30$, and we do not know the variance, then we have no good solution.

5 χ^2 Distribution

The χ^2 distribution measures the distribution of the sample variance, given a normal population.

6 Bayes

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$