Student's t-distribution

061601095 Toshihiko Terasawa
061601058 Naohiro Tsuzu

• Motivation.

Let $X_1, \ldots, X_n$ be i.i.d random variables with $X_j \sim N(\mu, \sigma^2)$

Our situation ; μ is known and σ is unknown.

Recall : $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$; the sample mean.

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$; the sample variance.

Then the random variable $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$

Now we consider the random variable which replaces σ with S in Z.

$$T = \frac{\overline{X} - \mu}{S / \sqrt{n}},$$

This random variable has a t-distribution with $n-1$ degrees of freedom.

• Basic properties of t-distribution

$t(\nu)$: t-distribution with ν degrees of freedom. ($\nu \in \mathbb{Z}_{>0}$)

- p.d.f : $f(x) = \dfrac{1}{\sqrt{\nu}\ B(\frac{\nu}{2}, \frac{1}{2})} \left(1 + \dfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$ where B is Beta function.

- mean $\quad 0$ for $\nu > 1$.

- variance $\quad \dfrac{\nu}{\nu - 2}$ for $\nu > 2$.

If n is larger than 30, then we can regard t-distribution as standard normal distribution approximately, so t-distribution is used when sample size is small (< 30).

• Application

In statistics, t-distribution is used for t-test.

For example, the heights of a random sample of basketball players could be compared with the heights from a random sample of football players.

The student's t distribution would be used to test whether the data indicated that one grup was significantly taller than the other. More precisely, it would be testing the hypothesis that both samples were drawn from the same normal population. A significant value of t would cause the hypothesis to be rejected, indicating that the means were significantly diffent.

Reference

- Catherine Forbes, Merran Evans, Nicholas Hastings, Brian Peacock, "Statistical Distributions"
- George Casella, Roger L. Berger "Statistical Inference"