T-test

Chao Cheng

August 2, 2022

1 Basic knowledge

$\phi(x)$ and $\Phi(x)$ are pdf and cdf of standard normal distribution, respectively. We use Z to represent a random variable that follows standard normal distribution and z_α the lower α quantile of standard normal distribution. Therefore

$$P(Z \le z_\alpha) = \Phi(z_\alpha) = \alpha.$$

Theorem 1. *Let $x_1, \cdots, x_n$ be a random sample from a population with mean μ and variance $\sigma^2 < \infty$. Then*

1. $\mathrm{E}\bar{x} = \mu$.
2. $\mathrm{Var}\bar{x} = \sigma^2/n$.
3. $\mathrm{E}S^2 = \sigma^2$, *where* $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

Theorem 2. *Let $x_1, \cdots, x_n$ be a random sample from $N(\mu, \sigma^2)$. Then*

1. $\bar{X} \sim N(\mu, \sigma^2/n)$.
2. $\bar{X}$ *is independent of* S^2.
3. $(n-1)S^2/\sigma^2$ *follows a chi-squared distribution with* $n-1$ *degree of freedom.*

2 One-sample test

Consider a random sample $x_1, \cdots, x_n$ from $N(\mu, \sigma^2)$. The likelihood is

$$\begin{aligned} f(x_1, \cdots, x_n) &= \prod_{i=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) \\ &= \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right). \end{aligned}$$

We propose the test

$$H_0:\ \mu = \mu_0 \quad \textbf{v.s} \quad H_1:\ \mu \ne \mu_0$$

2.1 variance known

Construct LRT

$$LR = \frac{\max_{\mu\in H_0} f\left(x_1,\cdots,x_n|\mu\right)}{\max_{\mu\in H_0\cup H_1} f\left(x_1,\cdots,x_n|\mu\right)} = \frac{f\left(x_1,\cdots,x_n|\mu=\mu_0\right)}{f\left(x_1,\cdots,x_n|\mu=\bar{x}\right)} = \exp\left(-\frac{\left(\bar{x}-\mu_0\right)^2}{2\sigma^2/n}\right)$$

Therefore rejecting H_0 when LR is smaller than some constant C is equivalent to rejecting H_0 when $|\bar{x}-\mu_0|$ is larger than some other constant C. Hence

$$\text{Reject Region: } \{\bar{x}: \quad |\bar{x}-\mu_0| > C\}$$

2.1.1 Decide C from α

From definition of α we know that C in the reject region is chosen such that

$$P\left(|\bar{x}-\mu_0| > C|H_0 \text{ is true }\right) \leq \alpha.$$

But to fully utilize the test, we choose to use equal sign instead of $\leq$. Therefore

$$P\left(|\bar{x}-\mu_0| > C|\mu=\mu_0\right) = \alpha.$$

Note that $\bar{x} \sim N\left(\mu, \sigma^2/n\right)$. Then under the condition $\mu=\mu_0$,

$$\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}} \sim N\left(0,1\right).$$

Therefore we propose the reject region for H_0 being

$$\left|\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}}\right| \geq z_{1-\alpha/2}.$$

2.1.2 Power at given underlying μ

The power (the probability to reject H_0, when H_1 is true) of the proposed test procedure for any given underlying $\mu \neq \mu_0$ is computed as

$$\begin{aligned}
&P\left(\left|\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}}\right| \geq z_{1-\alpha/2}\right) \\
=&P\left(\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}} \leq z_{\alpha/2}\right) + P\left(\frac{\bar{x}-\mu_0}{\sqrt{\sigma^2/n}} \geq z_{1-\alpha/2}\right) \\
=&P\left(\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}} \leq z_{\alpha/2} + \frac{\mu_0-\mu}{\sqrt{\sigma^2/n}}\right) + P\left(\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}} \geq z_{1-\alpha/2} + \frac{\mu_0-\mu}{\sqrt{\sigma^2/n}}\right) \\
=&P\left(Z \leq z_{\alpha/2} + \frac{\mu_0-\mu}{\sqrt{\sigma^2/n}}\right) + P\left(Z \geq z_{1-\alpha/2} + \frac{\mu_0-\mu}{\sqrt{\sigma^2/n}}\right)
\end{aligned}$$

Here we use the fact that $\frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}} \sim N\left(0,1\right)$. W.l.o.g, assume that $\mu > \mu_0$, then

$$P\left(Z \leq z_{\alpha/2} + \frac{\mu_0-\mu}{\sqrt{\sigma^2/n}}\right)$$

would be really close to zero and

$$P\left(Z \geq z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sqrt{\sigma^2/n}}\right)$$

will offer most of the power. In order to guarantee a power of at least $1 - \beta$, we could simply set

$$P\left(Z \geq z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sqrt{\sigma^2/n}}\right) = 1 - \beta$$

2.2 variance unknown

3 Two sample test

3.1 Two-sample, variance known

3.2 Two-sample, variance unknown but equal

3.3 Two-sample, variance unknown and unequal