STAT 423/523 Statistical Methods for Engineers and Scientists

Lecture 5: Hypothesis Testing II

Chencheng Cai

Washington State University

From Z-test to T-test

We discussed Z-test in the last lecture. The Z-test is based on the Z-statistic:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

where σ is the population standard deviation assumed to be known.

From Z-test to T-test

We discussed Z-test in the last lecture. The Z-test is based on the Z-statistic:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

where σ is the population standard deviation assumed to be known.

In practice, we often do not know the population standard deviation.

T-Statistic

For a population with unknown standard deviation, we consider the **T-statistic**:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}},$$

where S is the sample standard deviation:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

T-Statistic

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}},$$

When the population is normally distributed, the T-statistic, under $H_0 : \mu = \mu_0$, follows a t-distribution with $n - 1$ degrees of freedom.

$$T \sim t_{n-1}$$

T-Statistic

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}},$$

When the population is normally distributed, the T-statistic, under $H_0 : \mu = \mu_0$, follows a t-distribution with $n-1$ degrees of freedom.

$$T \sim t_{n-1}$$

The test based on the T-statistic is called the **T-test**, which is similar to the Z-test.

Two-sided T-Test

Let $X_1, \ldots, X_n$ be a random sample from a normal population with unknown mean μ and unknown standard deviation. We want to test

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_a : \mu \neq \mu_0.$$

The test statistic is

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}.$$

Two-sided T-Test

Let $X_1, \ldots, X_n$ be a random sample from a normal population with unknown mean μ and unknown standard deviation. We want to test

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_a : \mu \neq \mu_0.$$

The test statistic is

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}.$$

The decision rule is

$$\text{reject null if } |T| > t_{\alpha/2, n-1},$$

where $t_{\alpha/2,n-1}$ is the $\alpha/2$ quantile of the t-distribution with $n-1$ degrees of freedom.

Example (Textbook 8.10)

After 24 hours of smoke abstinence, 20 smokers are asked how much time has elapsed during a 45-second period. The sample data is summarized as follows:

$$\bar{x} = 59.30, \quad s = 9.84$$

Example (Textbook 8.10)

After 24 hours of smoke abstinence, 20 smokers are asked how much time has elapsed during a 45-second period. The sample data is summarized as follows:

$$\bar{x} = 59.30, \quad s = 9.84$$

We consider test the following hypothesis with $\alpha = 0.05$:

$$H_0 : \mu = 45 \quad \text{v.s.} \quad H_a : \mu \neq 45.$$

Example (Textbook 8.10)

After 24 hours of smoke abstinence, 20 smokers are asked how much time has elapsed during a 45-second period. The sample data is summarized as follows:

$$\bar{x} = 59.30, \quad s = 9.84$$

We consider test the following hypothesis with $\alpha = 0.05$:

$$H_0 : \mu = 45 \quad \text{v.s.} \quad H_a : \mu \neq 45.$$

The test statistic is

$$t = \frac{\bar{x} - 45}{s/\sqrt{n}} = 6.50$$

The quantile of the t-distribution is $t_{0.025,19} = 2.093$.

Example (Textbook 8.10)

After 24 hours of smoke abstinence, 20 smokers are asked how much time has elapsed during a 45-second period. The sample data is summarized as follows:

$$\bar{x} = 59.30, \quad s = 9.84$$

We consider test the following hypothesis with $\alpha = 0.05$:

$$H_0 : \mu = 45 \quad \text{v.s.} \quad H_a : \mu \neq 45.$$

The test statistic is

$$t = \frac{\bar{x} - 45}{s/\sqrt{n}} = 6.50$$

The quantile of the t-distribution is $t_{0.025,19} = 2.093$. We reject the null hypothesis since $|t| = 6.50 > 2.093$.

P-value of Two-sided T-Test

Similar to the Z-test, the p-value of the two-sided T-test is

$$\text{p-value} = 2(1 - F_{t,n-1}(|T|)),$$

where $F_{t,n-1}(\cdot)$ is the cdf of the t-distribution with $n-1$ degrees of freedom.

P-value of Two-sided T-Test

Similar to the Z-test, the p-value of the two-sided T-test is

$$\text{p-value} = 2(1 - F_{t,n-1}(|T|)),$$

where $F_{t,n-1}(\cdot)$ is the cdf of the t-distribution with $n-1$ degrees of freedom.

We reject the null hypothesis if the p-value is less than the significance level α.

Power of Two-sided T-Test

Consider $\mu_1 \neq \mu_0$ in the alternative hypothesis. The power $(1 - \beta(\mu_1))$ of the two-sided T-test is

$$
\begin{aligned}
\text{Power} &= P(\text{reject} \mid H_1) \\
&= P(|T| > t_{\alpha/2,n-1} | \mu = \mu_1) \\
&= F_{t,n-1}\left(-t_{\alpha/2,n-1} + \frac{\mu_1 - \mu_0}{s/\sqrt{n}}\right) + F_{t,n-1}\left(-t_{\alpha/2,n-1} - \frac{\mu_1 - \mu_0}{s/\sqrt{n}}\right)
\end{aligned}
$$

Sample Size Determination

In order to have a desired power $1 - \beta^*$, we need to determine the minimum sample size n such that $1 - \beta(\mu_1) \geq 1 - \beta^*$.

Sample Size Determination

In order to have a desired power $1-\beta^*$, we need to determine the minimum sample size n such that $1-\beta(\mu_1) \geq 1-\beta^*$.

Similar to the two-sided Z-test, the approximate formula for the sample size is

$$n^* = \left\lceil \left(\frac{s(t_{\alpha/2,n-1} - t_{1-\beta^*,n-1})}{\mu_1 - \mu_0} \right)^2 \right\rceil$$

Summary of Two-sided T-test

- Hypothesis:

$$H_0: \ \mu = \mu_0 \quad \text{v.s.} \quad H_a: \ \mu \neq \mu_0$$

- Test statistic:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

- Rejection region:

$$\text{reject null if } |T| > t_{\alpha/2,n-1}$$

- P-value:

$$\text{p-value} = 2(1 - F_{t,n-1}(|T|))$$

- Power:

$$1 - \beta(\mu_1) = F_{t,n-1}\left(-t_{\alpha/2,n-1} - \frac{\mu_1 - \mu_0}{s/\sqrt{n}}\right) + F_{t,n-1}\left(-t_{\alpha/2,n-1} + \frac{\mu_1 - \mu_0}{s/\sqrt{n}}\right)$$

- Sample size determination: (approximation)

$$n^* = \left\lceil \left(\frac{s(t_{\alpha/2,n-1} - t_{1-\beta^*,n-1})}{\mu_1 - \mu_0}\right)^2 \right\rceil$$

One-sided T-Test (Summary)

- Hypothesis:

$$H_0: \ \mu = \mu_0 \quad \text{v.s.} \quad H_a: \ \mu > \mu_0$$

- Test statistic:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

- Rejection region:

$$\text{reject null if } T > t_{\alpha,n-1}$$

- P-value:

$$\text{p-value} = 1 - F_{t,n-1}(T)$$

- Power:

$$1 - \beta(\mu_1) = 1 - F_{t,n-1}\left(t_{\alpha/2,n-1} - \frac{\mu_1 - \mu_0}{s/\sqrt{n}}\right)$$

- Sample size determination:

$$n^* = \left\lceil \left(\frac{s(t_{\alpha,n-1} - t_{1-\beta^*,n-1})}{\mu_1 - \mu_0}\right)^2 \right\rceil$$

Example (Textbook 8.9)

The failure stree for 19 carbon nanofibers are measured in MPa. The sample data is summarized as follows:

$$\bar{x} = 562.68, \quad s = 180.874$$

The researchers want to test whether the average failure stress is greater than 500 MPa with $\alpha = 0.05$.

Example (Textbook 8.9)

The failure stree for 19 carbon nanofibers are measured in MPa. The sample data is summarized as follows:

$$\bar{x} = 562.68, \quad s = 180.874$$

The researchers want to test whether the average failure stress is greater than 500 MPa with $\alpha = 0.05$.

Hypothesis:

$$H_0 : \mu = 500 \quad \text{v.s.} \quad H_a : \mu > 500$$

Example (Textbook 8.9)

The failure stree for 19 carbon nanofibers are measured in MPa. The sample data is summarized as follows:

$$\bar{x} = 562.68, \quad s = 180.874$$

The researchers want to test whether the average failure stress is greater than 500 MPa with $\alpha = 0.05$.

Hypothesis:

$$H_0 : \mu = 500 \quad \text{v.s.} \quad H_a : \mu > 500$$

Test statistic:

$$T = \frac{\bar{x} - 500}{s/\sqrt{n}} = 1.51$$

Example (Textbook 8.9)

The failure stree for 19 carbon nanofibers are measured in MPa. The sample data is summarized as follows:

$$\bar{x} = 562.68, \quad s = 180.874$$

The researchers want to test whether the average failure stress is greater than 500 MPa with $\alpha = 0.05$.

Hypothesis:

$$H_0 : \mu = 500 \quad \text{v.s.} \quad H_a : \mu > 500$$

Test statistic:

$$T = \frac{\bar{x} - 500}{s/\sqrt{n}} = 1.51$$

The p-value is $1 - F_{t,18}(1.51) = 0.074$.

Example (Textbook 8.9)

The failure stree for 19 carbon nanofibers are measured in MPa. The sample data is summarized as follows:

$$\bar{x} = 562.68, \quad s = 180.874$$

The researchers want to test whether the average failure stress is greater than 500 MPa with $\alpha = 0.05$.

Hypothesis:

$$H_0 : \mu = 500 \quad \text{v.s.} \quad H_a : \mu > 500$$

Test statistic:

$$T = \frac{\bar{x} - 500}{s/\sqrt{n}} = 1.51$$

The p-value is $1 - F_{t,18}(1.51) = 0.074$.
We failed to reject the null hypothesis since the p-value is greater than $\alpha = 0.05$.