STA2002: Probability and Statistics II

Hypothesis Testing II

Fangda Song and Ka Wai Tseng

School of Data Science, CUHK(SZ)

October, 2025

Introduction

In this lecture, we will

- Continue our journey in hypothesis testing
- Discuss hypothesis testing for:
  - Equality of two means
  - Equality of two variances
- Hypothesis testing involves:
  - Pooled t-test
  - Welch's t-test
  - Paired t-test
  - F-test
- Suggested reading: Chapter 8.1, 8.2 of the textbook

Recap of Hypothesis Testing Framework

1. Formulate H_0 and H_1, determine significance level α (e.g., 0.05 or 0.01)
2. Determine test statistic T, its distribution under H_0, and compute realized value t
3. Draw conclusion:
   - p-value approach: Reject H_0 if p-value $\leq \alpha$
   - Critical region approach: Reject H_0 if $t \in C$ (critical region)
   - Confidence interval approach: Reject H_0 if CI exclude null parameter value

Test Equality of Two Normal Means

- Two populations with i.i.d. samples:
  - $X_1, \ldots, X_n \sim N(\mu_X, \sigma_X^2)$
  - $Y_1, \ldots, Y_m \sim N(\mu_Y, \sigma_Y^2)$
- Test $H_0 : \mu_X = \mu_Y$ against some possible competing H_1

Three cases:

1. Pooled t-test: $X_1, \ldots, X_n$ and $Y_1, Y_2, \cdots, Y_n$ are independent, $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ unknown
2. Welch's t-test: $X_1, \ldots, X_n$ and $Y_1, Y_2, \cdots, Y_n$ are independent, $\sigma_X^2 \neq \sigma_Y^2$ unknown
3. Paired t-test: $n = m$, (X_i, Y_i) are independent pairs

Case 1: Pooled t-test

- Consider the pooled t-test statistic:

$$T := \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

where pooled variance:

$$S_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}$$

- Under H_0, $T \sim t(n+m-2)$
- Once we observe the data $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$, we can compute

$$t = \frac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}.$$

Case 1: Pooled t-test

For one-sided test $H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X > \mu_Y$:

- p-value is

$$p = \Pr(T > t; H_0) = \Pr(T > t; T \sim t(n+m-2))$$

- Decision rules: Reject H_0 at significance level α when

$$\begin{aligned}
& p = \Pr(T > t; T \sim t(n+m-2)) \leq \alpha \\
\Leftrightarrow & t \geq t_\alpha(n+m-2) \\
\Leftrightarrow & 0 \notin \left(\bar{x} - \bar{y} - t_\alpha(n+m-2)s_p\sqrt{\frac{1}{n} + \frac{1}{m}}, \infty\right)
\end{aligned}$$

Case 1: Pooled t-test

For one-sided test $H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X < \mu_Y$:

- p-value is

$$p = \Pr(T < t; H_0) = \Pr(T < t; T \sim t(n+m-2))$$

- Decision rules: Reject H_0 at significance level α when

$$\begin{aligned}
&p = \Pr(T < t; T \sim t(n+m-2)) \leq \alpha \\
\Leftrightarrow &t \leq -t_\alpha(n+m-2) \\
\Leftrightarrow &0 \notin (-\infty, \bar{x} - \bar{y} + t_\alpha(n+m-2)s_p\sqrt{\frac{1}{n} + \frac{1}{m}})
\end{aligned}$$

Case 1: Pooled t-test

For two-sided test $H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X \neq \mu_Y$:

- p-value is

$$p = \Pr(|T| > |t|; H_0) = 2\Pr(T > |t|; T \sim t(n+m-2))$$

- Decision rules: Reject H_0 at significance level α when

$$\begin{aligned} & p = 2\Pr(T > |t|; T \sim t(n+m-2)) \leq \alpha \\ \Leftrightarrow & |t| \geq t_{\alpha/2}(n+m-2) \\ \Leftrightarrow & 0 \notin (\bar{x} - \bar{y} - t_{\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}, \bar{x} - \bar{y} + t_{\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}) \end{aligned}$$

- **Note that:** Last interval is just the two-sided pooled t-interval

Example: Pooled t-test

Example

Suppose $X \sim N(\mu_X, \sigma^2)$ and $Y \sim N(\mu_Y, \sigma^2)$. We observe

- $n = 12$, $\bar{x} = 1076.75$, $s_X^2 = 29.30$
- $m = 12$, $\bar{y} = 1072.33$, $s_Y^2 = 26.24$

We would like to test

$$H_0 : \mu_X = \mu_Y \text{ vs } H_1 : \mu_X \neq \mu_Y$$

at significance level $\alpha = 0.10$.

- Compute test statistic:

$$t = \frac{1076.75 - 1072.33}{\sqrt{\frac{11(29.30)+11(26.24)}{22}\left(\frac{1}{12}+\frac{1}{12}\right)}} = 2.05$$

- Comparison: $|t| = 2.05 > t_{0.05}(22) = 1.717$ or $p = 2\Pr(T > 2.05; T \sim t(22)) = 0.0525 < 0.1$
- Decision: Reject H_0 at 10% significance level

- Consider the Welch's t-test statistic:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

- Under H_0, $T \sim t(r)$ approximately with

$$r = \left\lfloor \frac{\left(\frac{S_X^2}{n} + \frac{S_Y^2}{m}\right)^2}{\frac{1}{n-1}\left(\frac{S_X^2}{n}\right)^2 + \frac{1}{m-1}\left(\frac{S_Y^2}{m}\right)^2} \right\rfloor$$

- Once we observe the data $x_1, ..., x_n$ and $y_1, ..., y_m$, we can compute,

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}}.$$

- Here, only the two-sided hypothesis testing is discussed.

Case 2: Welch's t-test

For two-sided test $H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X \neq \mu_Y$:

- p-value is

$$p = \Pr(|T| > |t|; H_0) = 2\Pr(T > |t|; T \sim t(r))$$

- Decision rules: Reject H_0 at significance level α when

$$\begin{aligned} & p = 2\Pr(T > |t|; T \sim t(r)) \leq \alpha \\ \Leftrightarrow & |t| \geq t_{\alpha/2}(r) \\ \Leftrightarrow & 0 \notin \left(\bar{x} - \bar{y} - t_{\alpha/2}(r)\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}, \right. \\ & \left. \bar{x} - \bar{y} + t_{\alpha/2}(r)\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}} \right) \end{aligned}$$

- **Note that:** Last interval is just the Welch's t-interval

Case 3: Paired t-test

- Consider the following two-sided test,

$$H_0 : \mu_X = \mu_Y, \quad H_1 : \mu_X \neq \mu_Y$$

- Introduce D_i, $i = 1, 2, ..., n$,

$$D_i := X_i - Y_i$$

- We consider the paired statistic,

$$T := \frac{\bar{D}}{S_D/\sqrt{n}} = \frac{\bar{X} - \bar{Y}}{S_D/\sqrt{n}}.$$

- Under $H_0 : \mu_X = \mu_Y$, $T \sim t(n-1)$
- Here, only the two-sided hypothesis testing is discussed.

Case 3: Paired t-test

- Once we observe the data $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, we can compute
$$t = \frac{\bar{x} - \bar{y}}{s_D/\sqrt{n}}.$$
- p-value is
$$p = \Pr(|T| > |t|; H_0) = 2\Pr(T > |t|; T \sim t(n-1))$$
- Decision rules: Reject H_0 at significance level α when
$$\begin{aligned}
& p = 2\Pr(T > |t|; T \sim t(n-1)) \le \alpha \\
\Leftrightarrow & |t| \ge t_{\alpha/2}(n-1) \\
\Leftrightarrow & 0 \notin \left(\bar{x} - \bar{y} - t_{\alpha/2}(n-1)\frac{s_D}{\sqrt{n}}, \right. \\
& \left. \bar{x} - \bar{y} + t_{\alpha/2}(n-1)\frac{s_D}{\sqrt{n}} \right)
\end{aligned}$$
- **Note that:** Last interval is just the paired t-interval

Example: Paired t-test

Example

Twenty-four girls in the 9th and 10th grades were put on an ultraheavy rope-jumping program. Someone thought that such a program would increase their speed in the 40-yard dash. Let D equal the difference in time to run the 40-yard dash—the "before-program time" (X) minus the "after-program time." (Y) Assume that the distribution of D is approximately $N(\mu_d, \sigma_d^2)$. We shall test the null hypothesis $H_0 : \mu_D = 0$ against the alternative hypothesis $H_1 : \mu_D > 0$ at significance level $\alpha = 0.10$. We observe $n = 24$, $\bar{d} = 0.0788$, $s_D = 0.2549$.

Example: Paired t-test

- p-value approach:
  - $p = \Pr(T > 1.514; T \sim t(23)) = 0.0718 < 0.1$, so we reject H_0 at 10% significance level.
- Critical region approach:
  - $|t| = 1.514 > 1.319$, so we reject H_0 at 10% significance level.
- Confidence interval approach: The 90% one-sided paired CI is

$$(0.0788 - 1.319 \times \frac{0.2549}{\sqrt{24}}, \infty) = (0.0101, \infty),$$

  which doesn't include 0, so we reject H_0 at 10% significance level.

Test Equality of Two Normal Variances

- Suppose we have two populations with i.i.d. samples:
  - $X_1, \dots, X_n \sim N(\mu_X, \sigma_X^2)$
  - $Y_1, \dots, Y_m \sim N(\mu_Y, \sigma_Y^2)$

  and these two populations are independent.
- We would like to test:

$$H_0 : \sigma_X^2 = \sigma_Y^2 \text{ vs } H_1 : \sigma_X^2 \neq \sigma_Y^2$$

- We have

$$U = \frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(n-1)$$

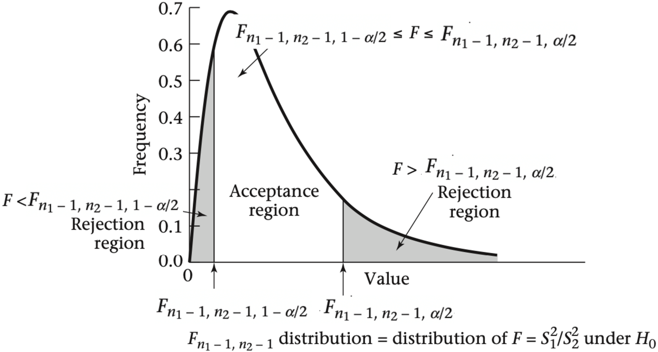
$$V = \frac{(m-1)S_Y^2}{\sigma_Y^2} \sim \chi^2(m-1)$$

Test Equality of Two Normal Variances

- Recall F-distribution definition:
  If $U \sim \chi^2(u)$, $V \sim \chi^2(v)$ independent, then:
  $$\frac{U/u}{V/v} \sim F(u, v)$$
- Define test statistic:
  $$F := \frac{U/(n-1)}{V/(m-1)} = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}$$
- Under H_0 ($\sigma_X^2 = \sigma_Y^2$), the ratio becomes:
  $$F = \frac{S_X^2}{S_Y^2} \sim F(n-1, m-1)$$
- Once we have the data $x_1, x_2, \cdots, x_n$ and $y_1, y_2, \cdots, y_m$, compute:
  $$f = \frac{s_X^2}{s_Y^2}$$

Test Equality of Two Normal Variances

- Intuitively, the data reject H_0 and favors H_1 when the ratio t is either too small or too large.



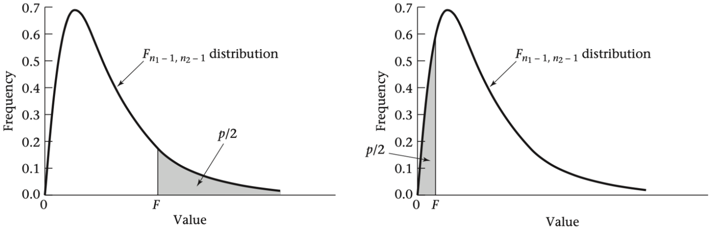
- Reject H_0 at significance level α when:

$$f \geq F_{\alpha/2}(n-1, m-1) \text{ or } f \leq F_{1-\alpha/2}(n-1, m-1)$$

- **Remark**: For F distribution, $F_{1-\alpha/2}(n-1, m-1) = \frac{1}{F_{\alpha/2}(m-1, n-1)}$

Test Equality of Two Normal Variances

For two-sided test $H_0 : \sigma_X^2 = \sigma_Y^2$ vs $H_1 : \sigma_X^2 \neq \sigma_Y^2$:



If $F = s_1^2/s_2^2 \geq 1$, then $p = 2 \times$ (area to the right of F under an $F_{n_1-1,\,n_2-1}$ distribution)

If $F = s_1^2/s_2^2 < 1$, then $p = 2 \times$ (area to the left of F under an $F_{n_1-1,\,n_2-1}$ distribution)

- When $f > 1$, p-value is

$$p = 2\Pr(F > f; F \sim F(n-1, m-1))$$

- When $f < 1$, p-value is

$$p = 2\Pr(F < f; F \sim F(n-1, m-1))$$

Test Equality of Two Normal Variances

Decision rules: Reject H_0 at significance level α when

$$
\begin{aligned}
& p = 2 \cdot \min\left\{\Pr(F > f), \Pr(F < f)\right\} \leq \alpha \\
\Leftrightarrow & f \geq F_{\alpha/2}(n-1, m-1) \quad \text{or} \quad f \leq F_{1-\alpha/2}(n-1, m-1) \\
\Leftrightarrow & 1 \notin \left(\frac{s_Y^2}{s_X^2} F_{1-\alpha/2}(n-1, m-1), \frac{s_Y^2}{s_X^2} F_{\alpha/2}(n-1, m-1) \right)
\end{aligned}
$$

Note: $\left[\frac{s_Y^2}{s_X^2} F_{1-\alpha/2}(n-1, m-1), \frac{s_Y^2}{s_X^2} F_{\alpha/2}(n-1, m-1)\right]$ is the two-sided $100(1-\alpha)\%$ CI of $\frac{\sigma_Y^2}{\sigma_X^2}$

Summary

H_0	H_1	**p-value**	**Critical Region**
Pooled t-test $(T \sim t(n+m-2),\ t = (\bar{x}-\bar{y})/(s_p\sqrt{\frac{1}{n}+\frac{1}{m}}))$			
$\mu_X = \mu_Y$	$\mu_X > \mu_Y$	$\Pr(T > t)$	$t \geq t_\alpha(n+m-2)$
$\mu_X = \mu_Y$	$\mu_X < \mu_Y$	$\Pr(T < t)$	$t \leq -t_\alpha(n+m-2)$
$\mu_X = \mu_Y$	$\mu_X \neq \mu_Y$	$2\Pr(T > \|t\|)$	$\|t\| \geq t_{\alpha/2}(n+m-2)$
Welch's t-test $(T \sim t(r),\ t = (\bar{x}-\bar{y})/(\sqrt{\frac{s_X^2}{n}+\frac{s_Y^2}{m}}))$			
$\mu_X = \mu_Y$	$\mu_X > \mu_Y$	$\Pr(T > t)$	$t \geq t_\alpha(r)$
$\mu_X = \mu_Y$	$\mu_X < \mu_Y$	$\Pr(T < t)$	$t \leq -t_\alpha(r)$
$\mu_X = \mu_Y$	$\mu_X \neq \mu_Y$	$2\Pr(T > \|t\|)$	$\|t\| \geq t_{\alpha/2}(r)$
Paired t-test $(T \sim t(n-1),\ t = \bar{d}/(s_D/\sqrt{n}))$			
$\mu_D = 0$	$\mu_D > 0$	$\Pr(T > t)$	$t \geq t_\alpha(n-1)$
$\mu_D = 0$	$\mu_D < 0$	$\Pr(T < t)$	$t \leq -t_\alpha(n-1)$
$\mu_D = 0$	$\mu_D \neq 0$	$2\Pr(T > \|t\|)$	$\|t\| \geq t_{\alpha/2}(n-1)$