# VE401 Recitation Class Note10

## Comparison Test

Chen Siyi
siyi.chen_chicy@sjtu.edu.cn

# 1 Comparison of Two Proportions

For large sample size:

$$\overline{X}^{(1)} \sim N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right), \quad \overline{X}^{(2)} \sim N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

So for large sample size:

$$\widehat{p_1 - p_2} = \widehat{p}_1 - \widehat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

Similarly we deduce the following $100(1-\alpha)\%$ confidence interval for $p_1 - p_2$:

$$\widehat{p}_1 - \widehat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}$$

## 1.1 Large-sample Test for Differences in Proportions

Suppose two random samples of (large) sizes $n_1$ and $n_2$ from two Bernoulli distributions with parameters $p_1$ and $p_2$ are given. Denote by $\widehat{p}_1$ and $\widehat{p}_2$ the means of the two samples.

Let $(\widehat{p}_1 - \widehat{p}_2)_0$ be a null value for the difference $p_1 - p_2$. Then the test based on the statistic

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - (p_1 - p_2)_0}{\sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}}$$

is called a large-sample test for differences in proportions.

We reject at significance level $\alpha$:

(i) $H_0 : p_1 - p_2 = (p_1 - p_2)_0$ if $|Z| > z_{\alpha/2}$

(ii) $H_0 : p_1 - p_2 \le (p_1 - p_2)_0$ if $Z > z_\alpha$

(iii) $H_0 : p_1 - p_2 \ge (p_1 - p_2)_0$ if $Z < -z_\alpha$

## 1.2 Pooled Test for Equality of Proportions

Suppose two random samples of (large) sizes $n_1$ and $n_2$ from two Bernoulli distributions with parameters $p_1$ and $p_2$ are given. Denote by $\widehat{p}_1$ and $\widehat{p}_2$ the means of the two samples.

Let $\widehat{p}$ be the pooled estimator for the proportion, which is defined as

$$\widehat{p} := \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1 + n_2}$$

Then the test based on the statistic

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

is called a pooled large-sample test for equality of proportions.

We reject at significance level $\alpha$:

(i) $H_0 : p_1 = p_2$ if $|Z| > z_{\alpha/2}$

(ii) $H_0 : p_1 \leq p_2$ if $Z > z_\alpha$

(iii) $H_0 : p_1 \geq p_2$ if $Z < -z_\alpha$

# 2 Comparison of Two Variances

## 2.1 The F-Distribution

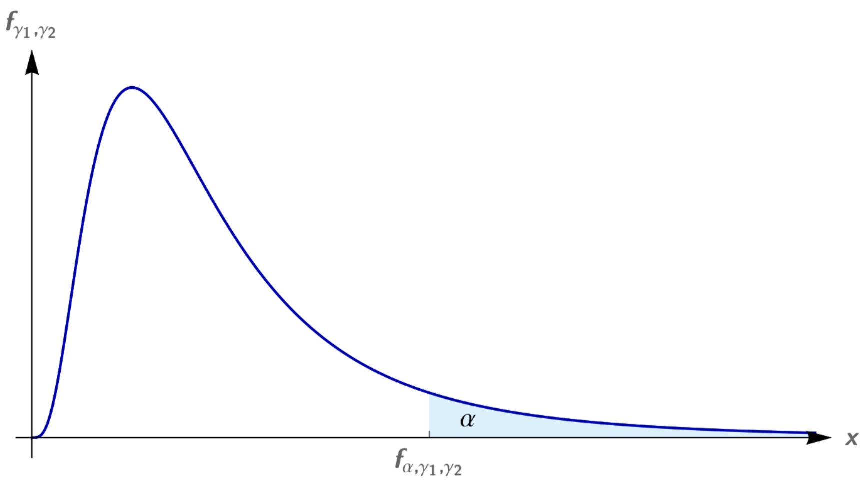
**Definition:**

Let $X^2_{\gamma_1}$ and $X^2_{\gamma_2}$ be independent chi-squared random variables with $\gamma_1$ and $\gamma_2$ degrees of freedom, respectively.

The random variable

$$F_{\gamma_1,\gamma_2} = \frac{X^2_{\gamma_1}/\gamma_1}{X^2_{\gamma_2}/\gamma_2}$$

is said to follow an F-distribution with $\gamma_1$ and $\gamma_2$ degrees of freedom.




**Properties:**

$$P\left[F_{\gamma_1,\gamma_2} < x\right] = P\left[\frac{1}{F_{\gamma_1,\gamma_2}} > \frac{1}{x}\right] = 1 - P\left[F_{\gamma_2,\gamma_1} < \frac{1}{x}\right]$$

$$f_{1-\alpha,\gamma_1,\gamma_2} = \frac{1}{f_{\alpha,\gamma_2,\gamma_1}}$$

## 2.2 The F-Test

**Statistic:**

Two Normally-Distributed Populations:

$$X^{(1)} \sim N\left(\mu_1, \sigma_1^2\right)$$
$$X^{(2)} \sim N\left(\mu_2, \sigma_2^2\right)$$

Taking samples of sizes $n_1$ and $n_2$ from the populations, we know that

$$\frac{(n_1-1)\,S_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}, \qquad \frac{(n_2-1)\,S_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$$

If $\sigma_1^2 = \sigma_2^2$, the statistic

$$S_1^2/S_2^2$$

follows an F -distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

**F-test:**

Let $S_1^2$ and $S_2^2$ be sample variances based on independent random samples of sizes $n_1$ and $n_2$ drawn from normal populations with means 1 and 2 and variances $\sigma_1^2$ and $\sigma_2^2$, respectively. Then a test based on the statistic

$$F_{n_1-1,n_2-1} = \frac{S_1^2}{S_2^2}$$

is called an F-test.

We reject at significance level $\alpha$:

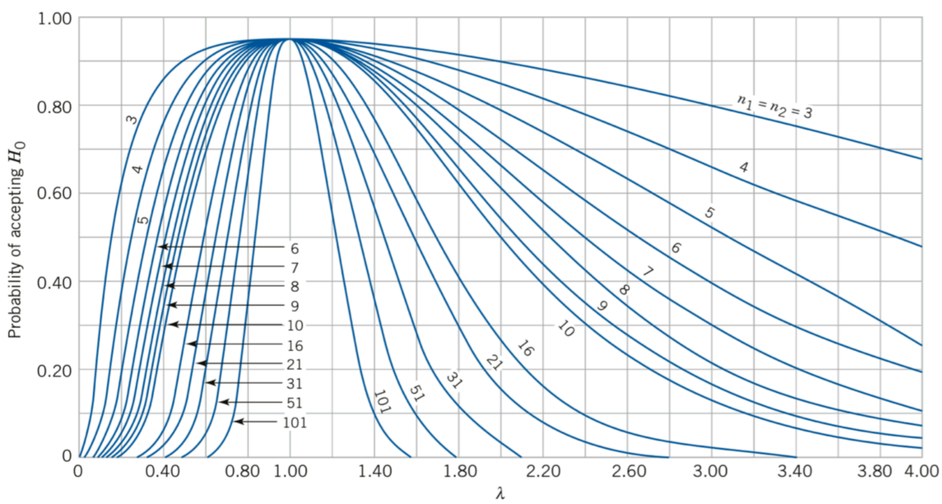
(i) $H_0 : \sigma_1 \leq \sigma_2$ if $\frac{S_1^2}{S_2^2} > f_{\alpha,n_1-1,n_2-1}$

(ii) $H_0 : \sigma_1 \geq \sigma_2$ if $\frac{S_2^2}{S_1^2} > f_{\alpha,n_2-1,n_1-1}$

(iii) $H_0 : \sigma_1 = \sigma_2$ if $\frac{S_1^2}{S_2^2} > f_{\alpha/2,n_1-1,n_2-1}$ or $\frac{S_2^2}{S_1^2} > f_{\alpha/2,n_2-1,n_1-1}$

**Abscissa of OC Curves (when $n_1 = n_2$):**

$$\lambda = \frac{\sigma_1}{\sigma_2}$$



**Comments:**

1. The populations must be normally distributed.
2. If possible, the sample sizes $n_1$ and $n_2$ should be equal.
3. The F-test is not very powerful, $\beta$ can be quite large.
4. We hope to not reject $H_0$

# 3 Comparison of Two Means

**Overview**

When comparing two means, what affect your choice of methods? Draw a map.

All four methods assume normality.

$$\overline{X}^{(1)} \sim N\left(\mu_1, \sigma_1^2/n_1\right), \quad \overline{X}^{(2)} \sim N\left(\mu_2, \sigma_2^2/n_2\right)$$

## 3.1 Variances Known

**Statistic:**
$\overline{X_1} - \overline{X_2}$ is normal with mean $\mu_1 - \mu_2$ and variance $\sigma_1^2/n_1 + \sigma_2^2/n_2$ So we can use the statistic to do Z-test:

$$Z = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

**Confidence Interval:**
100(1 - $\alpha$)% two sided confidence interval for $\mu_1 - \mu_2$

$$\mu_1 - \mu_2 = \overline{X}^{(1)} - \overline{X}^{(2)} \pm z_{\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

**Reject $H_0$:**
We reject at significance level $\alpha$:

(i) $\mathrm{H}_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $\left|\frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right| > z_{\alpha/2}$

(ii) $\mathrm{H}_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $\frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > z_\alpha$

(iii) $\mathrm{H}_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $\frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < -z_\alpha$

**Abscissa of OC Curves (when $n = n_1 = n_2$):**

$$d = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

**For $n_1 \neq n_2$:**
The table is used with the equivalent sample size

$$n = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

## 3.2 Equal but Unknown Variances

**Statistic:**

$$Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left(1/n_1 + 1/n_2\right)}}$$

Define the pooled estimator for the variance:

$$S_p^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}$$

Then the statistic:

$$T_{n_1+n_2-2} = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu_1 - \mu_2)}{\sqrt{S_p^2 \left(1/n_1 + 1/n_2\right)}}$$

**Confidence Interval:**

100(1 - $\alpha$)% two sided confidence interval for $\mu_1 - \mu_2$

$$\left(\overline{X}_1 - \overline{X}_2\right) \pm t_{\alpha/2, n_1+n_2-2} \sqrt{S_p^2 \left(1/n_1 + 1/n_2\right)}$$

**Reject $H_0$:**

We reject at significance level $\alpha$:

(i) $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|T_{n_1+n_2-2}| > t_{\alpha/2, n_1+n_2-2}$

(ii) $H_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} > t_{\alpha, n_1+n_2-2}$

(iii) $H_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} < -t_{\alpha, n_1+n_2-2}$

**Abscissa of OC Curves (when $n = n_1 = n_2$):**

$$d = \frac{|\mu_1 - \mu_2|}{2\sigma}$$

Remember when reading the OC curve, use a modified $n^* = 2n - 1$.

As before, when $\sigma$ is unknown, we must either use an estimate or express the deviation in terms of $\sigma$.

## 3.3 (Inequal) and Unknown Variances

**The Welch-Satterthwaite Approximation:**

Let $X^(1), \dots, X^(1)$ be k independent normally distributed random variables with variances $\sigma_1^2, \dots, \sigma_k^2$.

Let $s_1^2, \dots, s_k^2$ be sample variances based on samples of sizes $n_1, \dots, n_k$ from the k populations, respectively. Let $\lambda_1, \dots, \lambda_k > 0$ be positive real numbers and define

$$\gamma := \frac{\left(\lambda_1 s_1^2 + \cdots + \lambda_k s_k^2\right)^2}{\sum_{i=1}^{k} \frac{\left(\lambda_i s_i^2\right)^2}{n_i - 1}}$$

Then the following is approximately a chi-squared distribution with $\gamma$ degrees of freedom:

$$\gamma \cdot \frac{\lambda_1 s_1^2 + \lambda_2 s_2^2 + \cdots + \lambda_k s_k^2}{\lambda_1 \sigma_1^2 + \lambda_2 \sigma_2^2 + \cdots + \lambda_k \sigma_k^2}$$

**For the case k = 2, $\lambda_1 = 1/n_1$ and $\lambda_2 = n_1$**

$$\gamma = \frac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\frac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(S_2^2/n_2\right)^2}{n_2 - 1}}$$

The following is approximately a chi-squared distribution with $\gamma$ degrees of freedom:

$$\gamma \cdot \frac{S_1^2/n_1 + S_2^2/n_2}{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

And the statistic below follows a T-distribution with $\gamma$ degrees of freedom.

$$T_\gamma = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

**Welch's (pooled) T-test for Equality of Means**

$$T_\gamma = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

We reject at significance level $\alpha$:

(i) $H_0 : \mu_1 - \mu_2 = \left(\mu_1 - \mu_2\right)_0$ if $|T_\gamma| > t_{\alpha/2,\gamma}$

(ii) $H_0 : \mu_1 - \mu_2 \leq \left(\mu_1 - \mu_2\right)_0$ if $T_\gamma > t_{\alpha,\gamma}$

(iii) $H_0 : \mu_1 - \mu_2 \geq \left(\mu_1 - \mu_2\right)_0$ if $T_\gamma < -t_{\alpha,\gamma}$

**Comments**

1. In practice, we round $\gamma$ down to the nearest integer.
2. Power calculations are much more difficult. There are no simple OC curves for Welch's test.
3. As remarked earlier, it is not a good idea to pre-test for equal variances and then make a decision whether to use Student's or Welch's test. In fact, current recommendations are to always use Welch's test.

## 3.4 Paired T-Test

**Statistic**

Assume that X and Y follow a joint bivariate normal distribution. Then D = X - Y follows a normal distribution. Then:

$$T_{n-1} = \frac{\overline{D} - \mu_D}{\sqrt{S_D^2/n}}$$

**Reject $H_0$**

We reject at significance level $\alpha$:

(i) $\mathrm{H}_0 : \mu_\mathrm{D} = (\mu_\mathrm{D})_0$ if $|T_{n-1}| > t_{\alpha/2,n-1}$

(ii) $\mathrm{H}_0 : \mu_\mathrm{D} \leq (\mu_\mathrm{D})_0$ if $T_{n-1} > t_{\alpha,n-1}$

(iii) $\mathrm{H}_0 : \mu_\mathrm{D} \geq (\mu_\mathrm{D})_0$ if $T_{n-1} < -t_{\alpha,n-1}$

## 3.5 Paired vs. Pooled T-Tests

Positive relation makes a paired T-test more powerful.

1. $\rho_{XY} > 0$: Paired T-Test is more powerful
2. $\rho_{XY} \leq 0$: Pooled T-Test is more powerful

# 4 Test Correlation Coefficient

**Estimator:**

The natural unbiased estimators:

$$\widehat{\operatorname{Var}[X]} = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$
$$\operatorname{Cov}[X,Y] = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)$$

The natural choice for an estimator for the correlation coefficient is then:

$$R := \hat{\rho} = \frac{\sum \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum \left(X_i - \overline{X}\right)^2}\sqrt{\sum \left(Y_i - \overline{Y}\right)^2}}$$

**Statistic:**

When (X, Y) follows a bivariate normal distribution, and we set large sample size n, then the Fisher transformation of R

$$\frac{1}{2} \ln\left(\frac{1+R}{1-R}\right) = \operatorname{Artanh}(R)$$

is approximately normally distributed with

$$\mu = \frac{1}{2} \ln\left(\frac{1+\varrho}{1-\varrho}\right) = \operatorname{Artanh}(\varrho), \quad \sigma^2 = \frac{1}{n-3}$$

Therefore we have the statistic:

$$\begin{aligned} Z &= \frac{\sqrt{n-3}}{2}\left(\ln\left(\frac{1+R}{1-R}\right) - \ln\left(\frac{1+\varrho_0}{1-\varrho_0}\right)\right) \\ &= \sqrt{n-3}\left(\operatorname{Artanh}(R) - \operatorname{Artanh}\left(\varrho_0\right)\right) \end{aligned}$$

**$100(1-\alpha)\%$ Confidence Interval for $\rho$:**

$$\left[\frac{1+R-(1-R)e^{2z_{\alpha/2}/\sqrt{n-3}}}{1+R+(1-R)e^{2z_{\alpha/2}/\sqrt{n-3}}}, \frac{1+R-(1-R)e^{-2z_{\alpha/2}/\sqrt{n-3}}}{1+R+(1-R)e^{-2z_{\alpha/2}/\sqrt{n-3}}}\right]$$

$$\tanh\left(\operatorname{Artanh}(R) \pm \frac{z_{\alpha/2}}{\sqrt{n-3}}\right)$$

**Reject $H_0$:**

$H_0$: $\rho = \rho_0$ if $|\sqrt{n-3}\left(\operatorname{Artanh}(R) - \operatorname{Artanh}\left(\varrho_0\right)\right)| > z_{\alpha/2}$ or $\rho_0$ is not in the confidence interval

# 5 Non-Parametric Comparisons of Locations

## 5.1 The Wilcoxon Rank-Sum Test

**Statistic**

Let X and Y be two random samples following some continuous distributions.

Let $X_1$,...,$X_m$ and $Y_1$,...,$Y_n$, m $\leq$ n, be random samples from X and Y and associate the rank $R_i$, i = 1,...,m+n, to the $R_i^{th}$ smallest among the m + n total observations. If ties in the rank occur, the mean of the ranks is assigned to all equal values.

Then the test based on the statistic

$$W_m := \text{sum of the ranks of } X_1, \ldots, X_m$$

is called the Wilcoxon rank-sum test.

**Reject $H_0$ for small m,n**

We reject $H_0$ : P [X > Y ] = 1/2 (and similarly the analogous one-sided hypotheses) at significance level $\alpha$ if $W_m$ falls into the corresponding critical region.

**Reject $H_0$ for large m,n**

$W_m$ is approximately normally distributed with

$$\mathrm{E}\left[W_m\right] = \frac{m(m+n+1)}{2}, \quad \mathrm{Var}\left[W_m\right] = \frac{mn(m+n+1)}{12}$$

If there are many ties, the variance may be corrected by taking

$$\mathrm{Var}\left[W_m\right] = \frac{mn(m+n+1)}{12 - \sum_{\text{groups}} \frac{t^3+t}{12}}$$

Then

$$Z = \frac{W_m - E\left[W_m\right]}{\sqrt{\mathrm{Var}\left[W_m\right]}}$$

We reject at significance level $\alpha$:

(i) $\mathrm{H_0 : P\left[X_m > X_n\right] = 0.5}$ if $|\mathrm{Z}| > z_{\alpha/2}$

(ii) $\mathrm{H_0 : P\left[X_m > X_n\right] \leq 0.5}$ if $\mathrm{Z} > z_\alpha$

(iii) $\mathrm{H_0 : P\left[X_m > X_n\right] \geq 0.5}$ if $\mathrm{Z} < -z_\alpha$

## 5.2 Non-Parametric Paired Test

**Assumption**

Let X and Y be two independent random variables that follow the same distribution but differ only in their location, i.e., X' := X $-$ $\delta$ and Y are independent and identically distributed.

Then $\delta$ is the median of D = X $-$ Y, and D will be symmetric about $\delta$.

Notice, X and Y themselves do not need to be symmetric.

**Method**

Transformed to the Wilcoxon signed-rank test for median, for the random variable D.

Let's write out the transformation:

1. $\mathrm{H}_0 : M_{\mathrm{D}} = (M_{\mathrm{D}})_0 \longleftarrow$
2. $\mathrm{H}_0 : M_{\mathrm{D}} \leq (M_{\mathrm{D}})_0 \longleftarrow$
3. $\mathrm{H}_0 : M_{\mathrm{D}} \geq (M_{\mathrm{D}})_0 \longleftarrow$