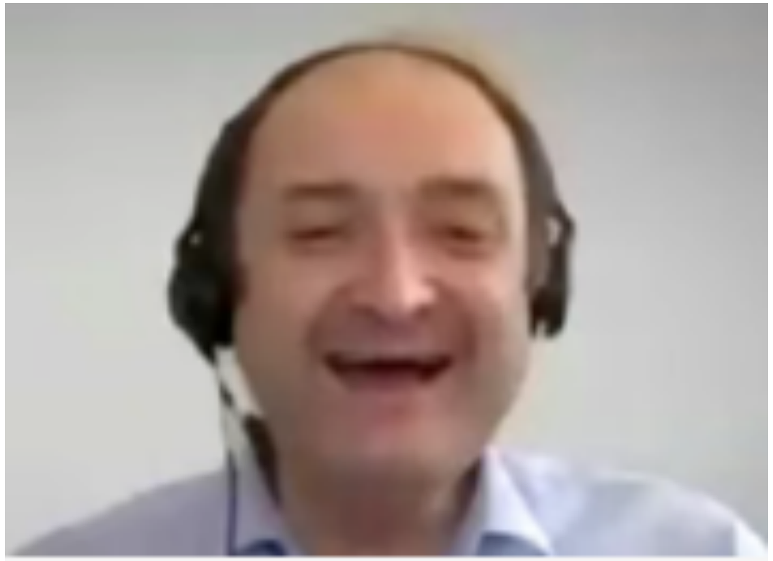
ECE4010J RC6

Shao Yujie



Test for Proportion

- When to use: we want to test the proportion of members of a population with some trait (or sample mean for n i.i.d. Bernoulli variables) p_0.
- What distribution does the testing statistics follow: Given a sample $X_1, \ldots, X_n$, $\frac{\overline{X}-p_0}{\sqrt{p_0(1-p_0)/n}}$ can be approximated by standard normal distribution Z (by Central Limit Theorem) for large sample size n.
- We reject at significance level α
  $H_0 : p = p_0$ if $|Z| > z_{\alpha/2}$,
  $H_0 : p \leq p_0$ if $Z > z_{\alpha}$,
  $H_0 : p \geq p_0$ if $Z < -z_{\alpha}$.

Test for Comparing Two Proportions

- When to use: we want compare two sample means for two groups (number:n_1, n_2) of Bernoulli variables (p_1, p_2) and test their differences $(p_1 - p_2)_0$
- What distribution does the testing statistics follow: Given two samples $X_{1_1}, \ldots, X_{1_n}$,$X_{2_1}, \ldots, X_{2_n}$, $\frac{\overline{X_1} - \overline{X_2} - (p_1 - p_2)}{\sqrt{\overline{X_1}(1-\overline{X_1})/n_1 + \overline{X_2}(1-\overline{X_2})/n_2}}$ can be approximated by standard normal distribution Z (by Central Limit Theorem) for large sample size n_1, n_2.
- We reject at significance level α
  $H_0 : p_1 - p_2 = (p_1 - p_2)_0$ if $|Z| > z_{\alpha/2}$,
  $H_0 : p_1 - p_2 \leq (p_1 - p_2)_0$ if $Z > z_\alpha$,
  $H_0 : p_1 - p_2 \geq (p_1 - p_2)_0$ if $Z < -z_\alpha$.
- Pooled test for Comparing Two Proportions: define **Pooled estimator** $\hat{p}$:$(n_1\hat{p}_1 + n_2\hat{p}_2)/(n_1 + n_2)$, the test statistics become: $\frac{\hat{P}}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$ (In pooled test, we assume $p_1 - p_2 = 0$)

F-Distribution

Definition

Let $\chi^2_{\gamma_1}$ and $\chi^2_{\gamma_2}$ be independent chi-squared random variables with γ_1 and γ_2 degrees of freedom, respectively. Then the random variable

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

follows a **F-distribution** with γ_1 and γ_2 degrees of freedom.

Property

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

Comparing Variances

Given two **normally-distributed** populations $X^{(1)} \sim N(\mu_1, \sigma_1^2)$ and $X^{(2)} \sim N(\mu_2, \sigma_2^2)$, we conduct an **F-test** to compare σ_1^2 and σ_2^2.

Procedure

Note that **normality** is required in comparison of variances.

1. Set up H_0 (and H_1 if applicable).
2. Set up **F-statistics** $F_{n_1-1,n_2-1} = \frac{S_1^2}{S_2^2}$.
3. Decide whether to reject H_0: we reject at significance level α
   - $H_0 : \sigma_1 \le \sigma_2$ if $S_1^2/S_2^2 > f_{\alpha,n_1-1,n_2-1}$
   - $H_0 : \sigma_1 \ge \sigma_2$ if $S_2^2/S_1^2 > f_{\alpha,n_2-1,n_1-1}$
   - $H_0 : \sigma_1 = \sigma_2$ if $S_1^2/S_2^2 > f_{\alpha/2,n_1-1,n_2-1}$ or $S_2^2/S_1^2 > f_{\alpha/2,n_2-1,n_1-1}$

OC Curve

For $n_1 = n_2 = n$, the abscissa is defined by $\lambda = \frac{\sigma_1}{\sigma_2}$.

Contents

Three Basic Cases

For two **normally distributed** populations:

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

When comparing μ_1 and μ_2, we consider three basic cases:

- σ_1^2 and σ_2^2 are known
- σ_1^2 and σ_2^2 are unknown but $\sigma_1^2 = \sigma_2^2$
- σ_1^2 and σ_2^2 are unknown and not necessarily equal

Case 1: Variance Known

Procedure

1. Set up H_0 (and H_1 if applicable).
2. Set up **Z-statistics** $Z = \frac{\overline{X}^{(1)} - \overline{X}^{(2)} - (\mu_1 - \mu_2)_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$.
3. Decide whether to reject H_0: we reject at significance level α
   - $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|Z| > z_{\alpha/2}$
   - $H_0 : \mu_1 - \mu_2 \le (\mu_1 - \mu_2)_0$ if $Z > z_\alpha$
   - $H_0 : \mu_1 - \mu_2 \ge (\mu_1 - \mu_2)_0$ if $Z < -z_\alpha$

OC Curve

- We can use the OC curves for normal distributions with $n = n_1 = n_2$ with $d = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$.
- When $n_1 \neq n_2$, we use the equivalent sample size $n = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2/n_1 + \sigma_2^2/n_2}$.

Case 2: Variance Equal but Unknown

Suppose we have random samples of sizes n_1, n_2 of $X^{(1)}$ and $X^{(2)}$, respectively and $X^{(1)} \sim N(\mu_1, \sigma^2), X^{(2)} \sim N(\mu_2, \sigma^2)$ with σ unknown.

- **Pooled estimator for variance**

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

- **T-statistics**

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

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

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

Case 2: Variance Equal but Unknown

Procedure

1. Set up H_0 (and H_1 if applicable).
2. Calculate the **pooled estimator for variance** $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$
3. Set up **T-statistics** $T_{n_1+n_2-2} = \frac{\overline{X}^{(1)}-\overline{X}^{(2)}-(\mu_1-\mu_2)_0}{\sqrt{S_p^2(1/n_1+1/n_2)}}$.
4. Decide whether to reject H_0: we reject at significance level α
   - $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|T_{n_1+n_2-2}| > t_{\alpha/2,n_1+n_2-2}$
   - $H_0 : \mu_1 - \mu_2 \leq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} > t_{\alpha,n_1+n_2-2}$
   - $H_0 : \mu_1 - \mu_2 \geq (\mu_1 - \mu_2)_0$ if $T_{n_1+n_2-2} < -t_{\alpha,n_1+n_2-2}$

Case 2: Variance Equal but Unknown

OC Curve

- We can use the OC curves for the T-test with $n = n_1 = n_2$ with $d = \frac{|\mu_1 - \mu_2|}{2\sigma}$
- We must use the modified sample size $n^* = 2n - 1$ when reading the charts.

*Note: When σ is unknown, we must either

- use an estimate (like S_p) or
- express the deviation in terms of σ

Case 3: Variance Not Necessarily Equal and Unknown

The distribution of $S_1^2/n_1 + S_2^2/n_2$ is unknown, so we need to use the **Welch-Satterthwaite approximation** to approximate the distribution and perform **Welch's T-Test**.

Procedure

1. Set up H_0 (and H_1 if applicable).
2. Calculate the **approximate degrees of freedom** $\gamma = \frac{(S_1^2/n_1+S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1}+\frac{(S_2^2/n_2)^2}{n_2-1}}$
3. Set up **T-statistics** $T_\gamma = \frac{\overline{X}^{(1)}-\overline{X}^{(2)}-(\mu_1-\mu_2)_0}{\sqrt{S_1^2/n_1+S_2^2/n_2}}$.
4. Decide whether to reject H_0: we reject at significance level α
   - $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$ if $|T_\gamma| > t_{\alpha/2,\gamma}$
   - $H_0 : \mu_1 - \mu_2 \le (\mu_1 - \mu_2)_0$ if $T_\gamma > t_{\alpha,\gamma}$
   - $H_0 : \mu_1 - \mu_2 \ge (\mu_1 - \mu_2)_0$ if $T_\gamma < -t_{\alpha,\gamma}$

Case 3: Variance Not Necessarily Equal and Unknown

Remarks

- The choice of γ. We round γ down to the nearest integer.
- The choice of test when σ_1^2 and σ_2^2 are unknown. We always use Welch's test. Welch's test is only slightly less powerful than Student's test even if the variances are equal; if they are unequal,Student's test is very unreliable.
- No simple OC curves for Welch's test.

Contents

Wilcoxon Rank-Sum Test

Let X and Y be two random samples following some **continuous distributions**. Decide whether to reject the null hypothesis

$$H_0 : P[X > Y] = \frac{1}{2} \quad \text{or} \quad H_0 : P[X > Y] \leq \frac{1}{2}$$

Procedure

1. Sort random samples from X, Y from smallest to largest (usually make $\#X < \#Y$) and assign rank R_i to the R_ith smallest one. If ties occur, assign the mean of the ranks to all the equal values.
2. Calculate the statistics $W_m = \sum R_i$ where R_i is the rank for X.
3. Reject or fail to reject at significance level α if W_m falls into the corresponding critical region.

Wilcoxon Rank-Sum Test

Rejection

We reject H_0 at significance level α

- For small m: check the table for critical values and compare the values with W_m.
- For large m ($m > 20$): perform a Z-test: W_m is approximately **normally distributed** with

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

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

$$\mathrm{Var}[W_m] = \frac{mn(m+n+1)}{12 - \sum \frac{t^3+t}{12}}$$

where t is the number of ties in each group of ties.

Example: Wilcoxon Rank-Sum Test

It has been suggested that the most highly motivated JI undergraduate students do at least as well, possibly even better, than graduate students in my graduate-level mathematics courses. For midterm exam scores of $Vv557$, the hypothesis is set as

$$H_0 : P[X_{\text{undergrad}} > X_{\text{grad}}] \leq 1/2$$

And we obtain the data as follows.

Graduate	5.5	5.5	12.75	18.75	19.25	11.25
	11.5	11.5	12.25	14.25	9.25	14.5
	13.25	8.25	16.75	10.5	6	15.25
	6.5	12.5	10.5	8.75	11.5	17
	2.75	13.25	19	16.5	11.5	1.75
Undergraduate	18.5	12.25	3	15	19.75	11.25
	11.75	19.25	12.25	19.75	16.25	13
	19.25	1.75				

Example: Wilcoxon Rank-Sum Test

Student	Points	Rank	Student	Points	Rank	Student	Points	Rank
grad	1.75	1.5	grad	11.5	17.5	undergrad	15	31
undergrad	1.75	1.5	grad	11.5	17.5	grad	15.25	32
grad	2.75	3	grad	11.5	17.5	undergrad	16.25	33
undergrad	3	4	grad	11.5	17.5	grad	16.5	34
grad	5.5	5.5	undergrad	11.75	20	grad	16.75	35
grad	5.5	5.5	grad	12.25	22	grad	17	36
grad	6	7	undergrad	12.25	22	undergrad	18.5	37
grad	6.5	8	undergrad	12.25	22	grad	18.75	38
grad	8.25	9	grad	12.5	24	grad	19	39
grad	8.75	10	grad	12.75	25	grad	19.25	41
grad	9.25	10	undergrad	13	26	undergrad	19.25	41
grad	10.5	12.5	grad	13.25	27.5	undergrad	19.25	41
grad	10.5	12.5	grad	13.25	27.5	undergrad	19.75	43.5
grad	11.25	14.5	grad	14.25	29	undergrad	19.75	43.5
undergrad	11.25	14.5	grad	14.5	30			

1. $W_m = 380$ (rank sum of undergrad scores)
2. Normal approximation with mean $\mathrm{E}[W_m] = \frac{m(m+n+1)}{2} = 315$ and $\mathrm{Var}[W_m] = \frac{mn(m+n+1)}{12} = 1575$. And $Z = \frac{W_m - \mathrm{E}[W_m]}{\sqrt{\mathrm{Var}[W_m]}} = 1.64$.
3. Find P-value: $P[Z \geq 1.64] = 0.0505$.

Paired T-Test

Condition and Methodology:

- Assumption: X and Y follow a **joint bivariate normal distribution** $\Rightarrow D = X - Y$ follows a **normal distribution**.
- Method: pair the samples as $D = X - Y$; perform a **paired T-test**.

Procedure

1. Set up the hypothesis, e.g., $H_0 : \mu_D \geq \mu_{D_0}$.
2. Pair X and Y and calculate $D = X - Y$.
3. Set up **T-statistics** $T_{n-1} = \frac{\overline{D}-\mu_D}{\sqrt{S_D^2/n}}$.
4. Reject or fail to reject by comparing the T-Statistics with critical values.

Non-Parametric Paired Test

Condition and Methodology:

- Assumption: X and Y do not follow a normal distribution, but **the same distribution** (differing in their location)
  $\Rightarrow D = X - Y$ is **symmetric** about its median.
- Method: pair the samples as $D = X - Y$; perform **Wilcoxon signed-rank test**.

Procedure

1. Set up the hypothesis, e.g., $H_0 : M_D \geq M_0$.
2. Pair X and Y and calculate $D = X - Y$.
3. Rank each $|D_i - M_0|$ and calculate W_+ or W_-.
4. Reject or fail to reject by comparing W_+ or W_- with critical values.

*Recall that the signed-rank test requires a random variable to have a symmetric distribution.

Paired vs. Pooled T-Tests

Assumption: X and Y are **normally distributed with equal variances** σ^2 and we take a paired sample of equal size n from (X, Y) to test the hypothesis

$$H_0 : \mu_X - \mu_Y = (\mu_X - \mu_Y)_0$$

We can perform a paired test or a pooled test.

Test Statistics

$$T_{\text{pooled}} = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)_0}{\sqrt{2S_p^2/n}}, \quad \text{critical value} = t_{\alpha/2,2n-2}$$

$$T_{\text{paired}} = \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)_0}{\sqrt{S_D^2/n}}, \quad \text{critical value} = t_{\alpha/2,n-1}$$

Paired vs. Pooled T-Tests

Compare the test statistics (denominators)

$$2S_p^2/n \text{ (estimates } 2\sigma^2/n) \quad \text{vs.} \quad S_D^2/n \text{ (estimates } \sigma_D^2/n)$$

by comparing $2\sigma^2/n$ with σ_D^2/n:

$$\begin{aligned}\frac{\sigma_D^2}{n} &= \frac{\mathrm{Var}[X-Y]}{n} = \frac{\mathrm{Var}[X]}{n} + \frac{\mathrm{Var}[Y]}{n} - \frac{2}{n}\mathrm{Cov}[X,Y] \\ &= \frac{2\sigma^2}{n} - \frac{2\sigma^2}{n}\frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}} = \frac{2\sigma^2}{n}(1-\rho_{XY})\end{aligned}$$

Conclusion

- $\rho_{XY} > 0$: paired T-test is more powerful since $T_{\text{paired}} > T_{\text{pooled}}$.
- $\rho_{XY} = 0$ (or < 0): pooled T-test is more powerful since it is easier to reject H_0 when comparing with $t_{\alpha/2,2n-2}$ than with $t_{\alpha/2,n-1}$.

$\Rightarrow$ **Positive correlation** makes a paired T-test more powerful.

Test for Correlation Coefficient

Estimating Correlation

The **estimator for correlation coefficient** R is

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

Distribution

Assumption: (X, Y) follows a **bivariate normal distribution**.

For large n, the Fisher transformation of R, $\frac{1}{2}\ln\left(\frac{1+R}{1-R}\right) = \operatorname{Artanh}(R)$ is approximately **normally distributed** with

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

Test for Correlation Coefficient

Assumption: (X, Y) follows a **bivariate normal distribution**.

- Hypothesis test. To test $H_0 : \varrho = \varrho_0$, we perform a **Z-test** on the Z-statistics

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

- Confidence interval. A $100(1-\alpha)\%$ confidence interval for ϱ

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

or

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