BUDT 730
Data, Models and Decisions

Lecture 8

Hypothesis Testing (II)

Prof. Sujin Kim

Agenda

- Study a hypothesis test for the population proportion; and
- Study hypothesis tests on comparing two means:
  - Paired samples
  - Two independent samples
- Also, study how to construct a CI for the difference between means

- Data Files:
  - `Customer Complaints.xlsx`
  - `Real_Estate.xlsx`
  - `BankTimes.xlsx`

Hypothesis Test

Population Proportion

$$H_0: p = p_0?$$

Hypothesis Tests for a Population Proportion

- To conduct a hypothesis test for a population proportion p, recall that the sampling distribution is approximately normal when the sample size, n is **sufficiently large** ($np \;\&\; n(1-p) \geq 5$)

$$\hat{p} \sim Normal\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

- Hypotheses: p_0 = Hypothesized population proportion
  - H_0: $p = p_0$ & H_1: $p \neq p_0$
  - H_0: $p \leq p_0$ & H_1: $p > p_0$
  - H_0: $p \geq p_0$ & H_1: $p < p_0$

Conducting the Hypothesis Test

- Test statistic for a proportion: Z-Test

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

- Critical values and rejection regions
  - H_1: $p \neq p_0$ → Reject H_0 if $|Z| > z_{1-\alpha/2}$
  - H_1: $p > p_0$ → Reject H_0 if $Z > z_{1-\alpha}$
  - H_1: $p < p_0$ → Reject H_0 if $Z < z_{\alpha}$

Example: Customer Complaints

- Walpole Appliance Company has a poor track record of responding promptly (within 30 days) to customers who submit complaints via mailed letters
- The department manager has set a goal of halving the proportion of late responses from 15% to 7.5% or less
- Sample data is collected from 400 cases, for which 23 were late
- Does it appear that the manager achieved his goal?
  1. What are the appropriate null and alternative hypothesis for this situation?
  2. Compute the critical values and p-value, and then make a decision for $\alpha = 5\%$.
  3. Compute the 95% CI for the proportion.

Example: Customer Complaints

What are the appropriate null and alternative hypothesis for this situation?

- H_0: $p \geq 0.075$
- H_1: $p < 0.075$

This is a one sample, one-tailed Z-test

Customer Complaints: “By Hand”

Compute test statistic:

- Given n = 400, 23 complaints had late responses

$$\hat{p} = \frac{23}{400} = 0.0575$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.0575 - 0.075}{\sqrt{\frac{0.075(1 - 0.075)}{400}}} = -1.329$$

- Critical value method:

- p-value method:

Customer Complaints: "By Hand"

Compute test statistic:

- Given n = 400, 23 complaints had late responses

$$\hat{p} = \frac{23}{400} = 0.0575$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.0575 - 0.075}{\sqrt{\frac{0.075(1 - 0.075)}{400}}} = -1.329$$

- Critical value method:

  $z_\alpha = z_{0.05}$ = NORM.S.INV(0.05) or qnorm(0.05) = -1.645

  Z=-1.329 > -1.645

- p-value method:

  p-value = NORM.S.DIST(-1.329,1) or pnorm(-1.329) = 0. 0920 > 0.05 (= α)

- Decision:

  Do not reject! The proportion of late responses is not statistically significantly less than 7.5%.

Customer Complaints: CI

$95\%\ CI = ?$

Customer Complaints: CI

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= 0.0575 \pm 1.96\sqrt{\frac{0.0575(1-0.0575)}{400}} = [0.035, 0.080]$$

Hypothesis test on comparing two means

Paired T-test and independent two sample T-test

Example: Sales Presentation Ratings

- Data file: `Real_Estate.xlsx`
- A real estate agent has collected a random sample of 75 houses that were recently sold in a suburban community. She is particularly interested in comparing the appraised value and recent selling price of the houses in this particular market (in $1,000).

House	Value	Price
1	119.37	121.87
2	148.93	150.25
⋮	⋮	⋮
75	141.78	139.35

Example: Sales Presentation Ratings

- **Objective**: Using this sample data, test weather there is a statistically significant mean difference between the appraised values and selling prices of the houses sold in this suburban community

Independent two samples vs. Paired samples

- For statistical reasons, we need to distinguish between paired samples and independent samples

  - **Paired samples – paired t-test**
    - Consider the difference between each pair as one single sample – one population
    - Perform a one-sample analysis on these differences to analyze the mean difference
    - Ex: husband-wife pairs
  - **Independent samples – (unpaired) two sample t-test**
    - Samples are from two independent population
    - Compare the two population means
    - Ex: Spending habits of men vs. women

Example: Sales Presentation Ratings

- Is the sample paired on not?
  - A paired-sample test is appropriate here because each value/price pair is for the same house, and they are **highly correlated**.
  - Note that two samples have the **same sample size**

Linear Correlation Table	**Value** **Real Estate**	**Price** **Real Estate**
Value	1.000	0.809
Price	0.809	1.000

Example: Sales Presentation Ratings

Set up the hypothesis test for this problem.

H_0:

H_1:

What test should we run?

( ) tailed, ( ) sample, ( ) test

Example: Sales Presentation Ratings

- Let μ_v = mean appraised value and μ_p= mean selling price
- We would like to estimate the difference between μ_v and μ_p.
- Let $d = \mu_v - \mu_p$
- Analyze the difference between the appraised value and the selling price as a single sample, then perform a one-sample analysis on the difference d.

House	Value	Price	Difference
1	119.37	121.87	-2.50
2	148.93	150.25	-1.32
⋮	⋮	⋮	
75	141.78	139.35	2.43

Example: Sales Presentation Ratings

- ***Two-tailed*** *hypotheses are* appropriate as the agent wants to simply compare the price, without assuming that one is greater than the other:
  - $H_0: \mu_v = \mu_p \quad or\ d = \mu_v - \mu_p = 0$
  - $H_1: \mu_v \neq \mu_p \quad or\ d = \mu_v - \mu_p \neq 0$
- We perform

  Two-tailed, paired sample, t-test

- Again, note that a paired t-test essentially performs a one-sample t-test on the difference $d = \mu_v - \mu_p$.

Perform the hypothesis test by hand.

What is your conclusion?

Example: Sales Presentation Ratings

- Sample size =75
- Sample mean = -0.3764
- Sample standard deviation = 9.1840

- Compute t-value and p-value.
- Make a decision for $\alpha = 1\%, 5\%$, and 10%, using two methods.

Example: Sales Presentation Ratings in R

> t.test(Real_Estate$Value, Real_Estate$Price, alternative="two.sided", paired = TRUE)

Paired t-test

Data: Real_Estate$Value and Real_Estate$Price

t = -0.35493 hours, df = 74, p-value = 0.7236

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.489448 1.736648

sample estimates:

mean of the differences

-0.3764

CI for Paired Samples

- We would like to estimate the difference between

$$\mu_v = \mathrm{E}(X_v) \text{ and } \mu_p = \mathrm{E}(X_p)$$

- Let $D = X_v - X_p$.
- CI for difference between means $d = (\mu_v - \mu_p)$:

$$\bar{D} \pm t - multiple * SE(\bar{D})$$

Example: Sales Presentation Ratings

- Sample size (n) =75
- Sample mean $(\bar{D})$ = - 0.3764
- Sample standard deviation (s) = 9.1840

- Compute 95% CI for the difference

Example: Customer's Waiting Time

- A bank has the business objective of improving the process for servicing customers during the noon-to-1PM lunch period
- To do so, the waiting time (defined as the number of minutes that elapses from when the customer enters the line until he or she reaches the teller window) needs to be shortened to increase customer satisfaction
- A random sample of customers at two different branches of the bank is selected and the waiting times are collected and stored in the file `BankTimes.xlsx`
- **Question:** Is there evidence that <u>the mean waiting time in Branch 1 is less than in Branch 2</u>?

Example: Customer’s Waiting Time

Is the sample paired on not?

Example: Customer's Waiting Time

- Is the sample paired on not?
  - (Unpaired) two-sample test is appropriate as samples are collected from two different branches and there is **almost no correlation** between them.

Linear Correlation Table	Branch1	Branch2
Branch1	1.000	0.022
Branch2	0.022	1.000

Example: Customer's Waiting Time

Set up the hypothesis test for this problem.

H_0:

H_1:

What test should we run?

(   ) tailed, (   ) sample, (   ) test

Example: Customer's Waiting Time

- One-tailed is appropriate as we want to prove that the mean waiting time in Branch 1 (μ_1) is less than the mean waiting time in Branch 2 (μ_2).

  $$H_0\text{: } \mu_1 \geq \mu_2$$

  $$H_1\text{: } \mu_1 < \mu_2$$

- *Thus, we conduct* ***one-tailed, two sample t-test.***

Perform the hypothesis test using R

What is your conclusion?

Equal-Variance Assumption

- The standard error of two-sample analysis depends on whether the standard deviations (or variances) from the two *populations* are equal or not
- How can you tell if they are equal?
  - Compare the sample standard deviations
  - StatTools reports the results of a statistical test (F-test) for **the equality of two population variances** in its Two-Sample output
  - If the **p-value is small**, say < 0.05, then they are probably **not equal**
- If they are equal, the sampling distribution for the difference between two **independent** sample means is the **t distribution** with $\mathbf{n_1 + n_2 - 2}$ degrees of freedom.
- What do you do if they are not equal? In this case, degree of freedom depends on n_1, n_2, s_1 and s_2.

Two-sample t-test for independent samples (you may skip this!)

- Equal- variance assumption:
  - Pooled (also known as combined or composite) standard deviation:

$$s_p = \sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}},$$

  - Standard error of difference between sample means:

$$SE(\overline{X_A} - \overline{X_B}) = s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

- Unequal-variance assumption:

$$SE(\overline{X_A} - \overline{X_B}) = \sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}$$

Example: Customer's Waiting Time in R

Equal- variance assumption:

```
> t.test(BankTimes$Branch1, BankTimes$Branch2, alternative="less", var.equal = TRUE)
        Two Sample t-test
data:  BankTimes$Branch1 and BankTimes$Branch2
```

t = -2.2485, df = 58, p-value = 0.01418 < 5%

```
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf -0.4018977
sample estimates:
mean of x mean of y
 5.239000  6.805333
```

Example: Customer's Waiting Time in R

Unequal-variance assumption:

```
> t.test(BankTimes$Branch1, BankTimes$Branch2, alternative="less", var.equal =FALSE )

        Welch Two Sample t-test
data:  BankTimes$Branch1 and BankTimes$Branch2
```

t = -2.2485, df = 52.911, p-value = 0.01437< 5%

```
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf -0.4000769
sample estimates:
mean of x mean of y
 5.239000  6.805333
```

Example: Customer's Waiting Time

- For both the equal variance and non-equal variance tests, p-value < 0.05
- We reject the null hypothesis with the significance level of 5%.
- Conclusion: **The mean waiting time in Branch 1 is less than the mean waiting time in Branch 2.**

Summary:
T-test for the comparison of two population means

- **Paired samples – paired t-test**
  - Two samples are **highly correlated**
  - The **sample sizes are the same**
  - Consider the differences between each pair as one single sample, and perform a **one-sample analysis** on these differences

- **Independent samples – two sample t-test**
  - Samples are from two independent population -two samples are almost **uncorrelated**
  - The sample sizes are not **necessarily the same**
  - To analyze the differences between the two samples, perform a **two-sample analysis** on the two samples

Next …

- Ch10 Regression Models