Lecture 22

Paired T-Tests

STAT 8010 Statistical Methods I
October 11, 2019

Whitney Huang
Clemson University

Example

A simple random sample with sample size 37 is taken and are subjected to a treatment ($\bar{X}_1 = 19.45, s_1 = 4.3$). A simple random sample with sample size 31 is taken and given a placebo ($\bar{X}_2 = 18.2, s_2 = 2.2$). At the 10% level can we say that the means are different between the two groups?

1. $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$
2. Need to decide between pooled or non-pooled procedure: $F_{obs} = \frac{4.3^2}{2.2^2} = 3.82 > F_{0.01}(36, 30) = 1.585 \Rightarrow$ Reject $\sigma_1 = \sigma_2$

Paired T-Tests

Motivating Example

Insurance handlers are concerned about the high estimates they are receiving for auto repairs from garage I compared to garage II. To verify their suspicions, each of 15 cars recently involved in an accident was taken to both garages for separate estimates of repair costs. The estimates from the two garages are given in the following table

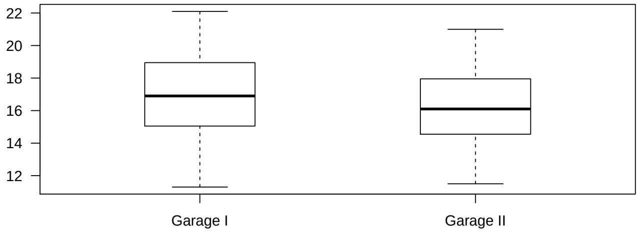
Garage I	Garage II	Garage I	Garage II	Garage I	Garage II
17.6	17.3	20.2	19.1	19.5	18.4
11.3	11.5	13.0	12.7	16.3	15.8
15.3	14.9	16.2	15.3	12.2	12.0
14.8	14.2	21.3	21.0	22.1	21.0
16.9	16.1	17.6	16.7	18.4	17.5

Example Cont'd

Sample statistics: $\bar{X}_1 = 16.85, \bar{X}_2 = 16.23, s_1 = 3.20, s_2 = 2.94$

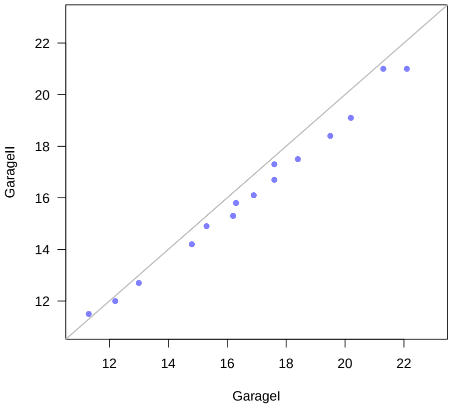
- $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 > 0$
- $t_{obs} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{16.85 - 16.23}{\sqrt{\frac{3.2^2}{15} + \frac{2.94^2}{15}}} = \frac{0.62}{1.12} = 0.55$
- Critical value for rejection region: $t_{0.05,df=27} = 1.70$
- Since t_{obs} is not in the rejection region. We fail to reject H_0 at 0.05 level.

Boxplots and R Output




```
        Welch Two Sample t-test

data:  GarageI and GarageII
t = 0.54616, df = 27.797, p-value =
0.2947
alternative hypothesis: true difference in means
is greater than 0
95 percent confidence interval:
 -1.29749      Inf
sample estimates:
mean of x mean of y
 16.84667  16.23333
```



For all but one of the 15 cars, the estimate from garage I was higher than that from garage II.

Analyzing Matched Pairs

- Matched pairs are dependent samples where each unit in the first sample is directly linked with a unit in the second sample
- This could occur in several situations, for example, before/after study, study on twins, pairing subjects based on similar characteristics
- We need different strategy for testing two dependent samples $\Rightarrow$ Paired T-Tests

Paired T-Tests

- $H_0 : \mu_{diff} = 0$ vs. $H_a : \mu_{diff} > 0$ (Upper-tailed); $\mu_{diff} < 0$ (Lower-tailed); $\mu_{diff} \neq 0$ (Two-tailed)
- Test statistic: $t^* = \frac{\bar{X}_{diff} - 0}{\frac{s_{diff}}{\sqrt{n}}}$. If $\mu_{diff} = 0$, then $t^* \sim t_{df=n-1}$
- Use rejection region method or P-value method to make a decision

Car Repair Example Revisited

Garage I - Garage II	Garage I - Garage II	Garage I - Garage II
17.6 - 17.3 = 0.3	20.2 - 19.1 = 1.1	19.5 - 18.4 = 1.1
11.3 - 11.5 = -0.2	13.0 - 12.7 = 0.3	16.3 - 15.8 = 0.5
15.3 - 14.9 = 0.4	16.2 - 15.3 = 0.9	12.2 - 12.0 = 0.2
14.8 - 14.2 = 0.6	21.3 - 21.0 = 0.3	22.1 - 21.0 = 1.1
16.9 - 16.1 = 0.8	17.6 - 16.7 = 0.9	18.4 - 17.5 = 0.9

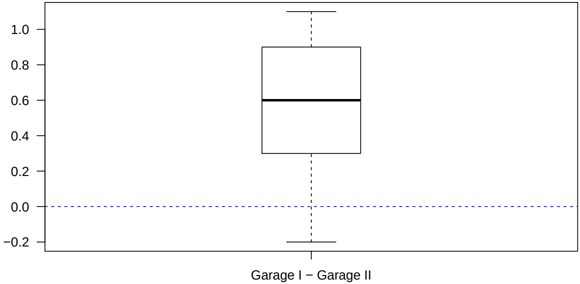
1. First, compute the difference in paired samples
2. Compute the sample mean and standard deviation for the differences
3. Then perform a one sample t-test

Car Repair Example Cont'd

$$\bar{X}_{diff} = 0.61, s_{diff} = 0.39$$

1. $H_0 : \mu_{diff} = 0$ vs. $H_a : \mu_{diff} > 0$
2. $t_{obs} = \frac{0.61}{\frac{0.39}{\sqrt{15}}} = 6.03$
3. Critical value for rejection region: $t_{0.05,df=14} = 1.76 \Rightarrow$ reject H_0
4. We do have enough evidence that the true mean repair cost difference for the garage I and II is greater than 0

Boxplot and R Output




```
        Paired t-test

data:  GarageI and GarageII
t = 6.0234, df = 14, p-value = 1.563e-05
alternative hypothesis: true difference in means
is greater than 0
95 percent confidence interval:
 0.4339886       Inf
sample estimates:
mean of the differences
              0.6133333
```

Summary

In this lecture, we learned

- Hypothesis Testing for $\mu_1 - \mu_2$ when $\sigma_1 \neq \sigma_2$
- Tests with matched samples

In next lecture we will learn

- Analysis of Variance (ANOVA)