

## **Lab 9: Comparing Two Means—Paired and Unpaired t-tests and Multiple Comparison Corrections**

### **Introduction and Objective:**

Before the second lecture exam, we learned how to calculate P values to help us make crisp decisions if we want to compare the means of two groups of values. The methods we learned are paired and unpaired t-test analyses.

We focused on the Student's t-test, which was developed by William Sealy Gosset, who published his research under his pseudonym of "Student". We also mentioned that if the variances of the two groups are unequal, reflected by very different standard deviations, we may consider using a modified t-test, Welch's t-test. How to calculate the P values using the Student's t-test was covered in our lecture and the second exam. The complicated Welch's t-test was mentioned, but how to calculate the P values by hand using Welch's t-test was not covered in our lecture. In fact, as we talked about it in the lecture, as the t-test is very robust, even the variances of the two groups of data are unequal, biologists/medical researchers may go ahead to calculate the pooled standard deviation for unpaired t-test anyway.

In today's lab, we will learn how to use R to calculate the P values using paired and unpaired t-test. You will also learn how to do the Welch's t-test. Besides, I will show you how to perform a t-test using Excel.

In addition, we are going to learn how to use R to do Bonferroni and FDR corrections of P values from multiple comparisons.

There is an in-class assignment due at the end of the lab.

The following files you should have received and are needed for today's class:

- 1) Midterms.csv
- 2) MouseMultipleComparisonsPCorrections.csv
- 3) MouseLifeSpan.csv
- 4) MouseMultipleComparisonsPCorrectionsAssignment.csv

All the R command lines are in **bold**; all the notes are following a #; all the R results directly follow the R codes/command lines and are not in bold or following a #.

```
setwd("C:/R")
```

## **I. Paired t-test:**

# The second midterm exam seems to be harder than the first one, and the average of the second exam is lower than the one from the first exam. Whether the difference between the two average scores is statistically significant? Let's use the scores of our two midterm exams to perform some statistical tests.

```
a<-read.csv("Midterms.csv")
```

```
View(a)
```

```
dim(a)
```

```
[1] 37 3
```

```
mid1<-a[,2]
```

```
mid2<-a[,3]
```

```
mid1m<-mean(mid1)
```

```
mid2m<-mean(mid2)
```

```
mid1m
```

```
[1] 82.39189
```

```
mid2m
```

```
[1] 79.36486
```

# Then, first, let's take a look at the variances of the data. They are very similar to each other:

```
sd(mid1)
```

```
[1] 9.330457
```

```
sd(mid2)
```

```
[1] 9.047443
```

# If you want to compute a P value to help you make a crisp decision, what kind of statistical test would you perform?

# Paired t-test! It is because the two sets of data are from the same objects (the same students).

# We set  $\alpha=0.05$

# Our null hypothesis ( $H_0$ ): there is no difference between the average scores from the two exams; any observed difference would result from pure coincidences or errors.

# We can use the R built-in function, `t.test()`, to perform the t-test:

**`t.test(mid1,mid2,paired=TRUE,var.equal=TRUE)`**

# In the brackets, the “mid1” and “mid2” are the two data sets you are going to compare. The “paired=TRUE” is telling R that you desire to do a paired t-test. The “var.equal=TRUE” means the variances (in fact, we checked the standard deviations) of the two data sets/samples (the 2 sets of the scores) are approximately the same or are not very different from each other.

# Once you hit “Enter” or “Return”, you may get the results as follows:

Paired t-test

data: mid1 and mid2

$t = 1.8266$ ,  $df = 36$ ,  $p\text{-value} = 0.07606$

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

95 percent confidence interval:

-0.3338778 6.3879319

sample estimates:

mean of the differences

3.027027

# From the results, you can see, R calculated the t-ratio, the degrees of freedom, the differences between the 2 means, the p-value (0.07606) as well as the confidence interval at 95% confidence level. R all showed you the alternative hypothesis.

# Since  $P > 0.05$ , we fail to reject our null hypothesis, and thus the difference between the average scores from the two exams is not statistically significant.

### # Unpaired t-test

# Although you should choose to do a paired t-test, what if you decided to do an unpaired t-test for this data anyway?

# You may type the command as follows:

**t.test(mid1,mid2,paired=FALSE,var.equal=TRUE)**

# As you can see, the only difference you made to the codes that you typed for doing a paired t-test is changing the “TRUE” to “FALSE” after the argument “paired=”. This informs R that you wish to perform an unpaired t-test.

# Once you hit “Enter” or “Return”, you may get the results as follows:

### Two Sample t-test

data: mid1 and mid2

t = 1.4167, df = 72, p-value = 0.1609

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

95 percent confidence interval:

-1.232289 7.286343

sample estimates:

mean of x mean of y

82.39189 79.36486

# As you can see, the p-value ( $P=0.1609$ ) you got is larger than the one ( $P=0.07606$ ) from using a paired t-test. As we talked about it in the lecture, a paired t-test tends to generate a smaller p-value.

**Challenge Question:** If you accidentally choose paired t-test instead of using an unpaired t-test, you may get more or less Type I errors? (The answer is at the end of this lab's notes.)

# Let's use another example we used in the lecture:

# To investigate if the infection control and aseptic care of Hospital A and Hospital B are at the same level, and neither is causing more nosocomial infection than the other, stool samples of patients from both hospitals were sent for testing the abundance of a type of bacteria, *Klebsiella pneumoniae*. The genomic DNA of this bacteria was quantified by qPCR. The delta CT values were reported as following:

Hospital A: 186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131, 149, 135, 132

Hospital B: 129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143, 152, 146, 144

# There are 20 samples from Hospital A, and only 17 from Hospital B. The sample numbers are uneven. So, apparently, we have to perform an unpaired t-test.

# First,  $H_0$ : There is no difference between the 2 mean delta CT values from the samples acquired from the 2 hospitals. Any observed difference results from coincidence or sampling/experimental errors.

# Second, set  $\alpha=0.05$

```
Hos1<-c(186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131, 149, 135, 132)
```

```
Hos2<-c(129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143, 152, 146, 144)
```

```
Hos1m<-mean(Hos1)
```

```
Hos2m<-mean(Hos2)
```

```
Hos1m
```

```
[1] 156.85
```

```
Hos2m
```

```
[1] 122.4706
```

```
sd(Hos1)
```

```
[1] 22.64201
```

```
sd(Hos2)
```

```
[1] 25.48313
```

# The two standard deviations are quite similar, we can go ahead to do a Student's unpaired t-test.

```
t.test(Hos1,Hos2,paired=FALSE, var.equal=TRUE)
```

### Two Sample t-test

data: Hos1 and Hos2

t = 4.3455, df = 35, p-value = 0.0001137

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

95 percent confidence interval:

18.31825 50.44057

sample estimates:

mean of x mean of y

156.8500 122.4706

# As you can see, the mean of x and mean of y equal to Hos1m and Hos2m, respectively.

# If you check the lecture PPT slide, you will see the t-statistic (t-ratio) we calculated by hand was 4.346, which equals the “t” we got here using R.

# Since  $P < 0.05$ , we reject our null hypothesis, and thus the difference of average bacteria numbers from the patients of the two hospitals is statistically significant. Since the mean from Hospital A is larger than the mean from Hospital B, Hospital A has worse aseptic care or infection control.

## II. Welch's t-test

# What if you think the two standard deviations are very different from each other (standard deviation of Hos2 is 25.48313, and you think it is much larger than 22.64201, which is the standard deviation of Hos1)?

# As we said, you may consider using the modified t-test, Welch's t-test. Although we did not discuss how to perform the Welch's t-test manually, as the calculations are bit complicated, we can perform it with ease in R. You can simply type the command as follows:

**t.test(Hos1,Hos2,paired=FALSE, var.equal=FALSE)**

# As you can see, the only change you made to the code was changing "TRUE" to "FALSE" after "var.equal=". This tells R, you think the variances of the two samples are very different. And then R would choose to do a Welch's t-test.

# Once you hit "Enter" or "Return", you may get the results as follows:

### Welch Two Sample t-test

data: Hos1 and Hos2

$t = 4.3031$ ,  $df = 32.394$ ,  $p\text{-value} = 0.0001455$

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

95 percent confidence interval:

18.11306 50.64577

sample estimates:

mean of x mean of y

156.8500 122.4706

# As you can see, if you choose to use Welch's t-test, the t-statistic (t-ratio), the df, and the P value are all different. You get a larger p-value, which reduces the chance to get a false-positive result.

### III. Use Microsoft Office Excel to perform a t-test:

# Click a cell, where you want to put your p-value:

Student	Midterm1	Midterm2
AE	81	81.5
AN	78.5	59
AS	82	83.5
CA	94	79
CF	77	69
DN	87	75
DR	99.5	77
EA	74	70
EC	75.5	83
EE	70.5	66
EI	79	92.5
EL	96.5	90
EP	81.5	74.5
ET	65.5	70
EU	93	92.5
HM	87.5	81.5
IA	89.5	88.5
AA	83.5	89.5
LA	90.5	71.5
LM	77	92.5
LW	86	78.5
NA	86	76.5
OA	81.5	80
DI	75	65.5

# In this cell, type “=TTEST(

Student	Midterm1	Midterm2
AE	81	81.5
AN	78.5	59
AS	82	83.5
CA	94	79
CF	77	69
DN	87	75
DR	99.5	77
EA	74	70
EC	75.5	83
EE	70.5	66
EI	79	92.5
EL	96.5	90
EP	81.5	74.5
ET	65.5	70
EU	93	92.5
HM	87.5	81.5
IA	89.5	88.5
AA	83.5	89.5
LA	90.5	71.5
LM	77	92.5
LW	86	78.5
NA	86	76.5
OA	81.5	80
DI	75	65.5

# Then, select the data you would like to compare, separate the 2 data sets using a comma,



	A	B	C	D	E	F
1	Student	Midterm I	Midterm II	Paired t-test		
2	AE	81	81.5	=TTEST(B2:B38,C2:C38)		
3	AN	78.5	59			
4	AS	82	83.5			
5	CA	94	79			
6	CF	77	69			
7	DN	87	75			
8	DR	99.5	77			
9	EA	74	70			
10	EC	75.5	83			
11	EE	70.5	66			
12	EI	79	92.5			
13	EL	96.5	90			
14	EP	81.5	74.5			
15	ET	65.5	70			
16	EU	93	92.5			
17	HM	87.5	81.5			
18	IA	89.5	88.5			
19	AA	83.5	89.5			
20	LA	90.5	71.5			
21	LM	77	92.5			
22	LW	86	78.5			
23	NA	86	76.5			
24	OA	81.5	80			

# Then, type another comma, followed by “2”. The “2” tells EXCEL to calculate a 2-tailed P value. If you type “1”, it will calculate a 1-tailed P value.

# And then, type a third comma, followed by “1” for the paired t-test. If you type “2”, it is for unpaired Student’s t-test with equal variance. If you type “3”, it will do Welch’s t-test for data with unequal variances.

File Home Insert Page Layout Formulas Data Review						
Clipboard		Font				
SUM =TTEST(B2:B38,C2:C38,2,1)						
	A	B	C	D	E	F
1	Student	MidtermI	MidtermII	Paired t-test		
2	AE	81	81.5	=TTEST(B2:B38,C2:C38,2,1)		
3	AN	78.5	59			
4	AS	82	83.5			
5	CA	94	79			
6	CF	77	69			
7	DN	87	75			
8	DR	99.5	77			
9	EA	74	70			
10	EC	75.5	83			
11	EE	70.5	66			
12	EI	79	92.5			
13	EL	96.5	90			
14	EP	81.5	74.5			
15	ET	65.5	70			
16	EU	93	92.5			
17	HM	87.5	81.5			
18	IA	89.5	88.5			
19	AA	83.5	89.5			
20	LA	90.5	71.5			
21	LM	77	92.5			
22	LW	86	78.5			
23	NA	86	76.5			
24	OA	81.5	80			
25	OI	75	65.5			

# Once you click “Enter”, you will see you get the P value as 0.076058362 (the same result as we got by using R).

#### IV. Multiple Comparison Corrections

**# Bonferroni Correction:**

```
MCC<-read.csv("MouseMultipleComparisonsPCorrections.csv")
```

```
View(MCC)
```

```
dim(MCC)
```

```
[1] 30 16
```

```
PS<-MCC[,16]
```

```
mean(PS)
```

```
[1] 5.492624e-07
```

```
PA<-p.adjust(PS,method="bonferroni",n=30)
```

# The argument method= is for which correction method you choose. The argument n= is for the number of comparisons you did.

```
PA
```

```
[1] 5.744970e-07 7.481970e-07 1.160988e-06 1.934868e-06 2.782716e-06
```

```
[6] 3.191760e-06 4.993470e-06 5.637930e-06 5.828880e-06 8.953350e-06
```

```
[11] 8.959530e-06 1.070571e-05 1.120008e-05 1.176789e-05 1.242369e-05
```

```
[16] 1.640547e-05 1.677387e-05 1.744896e-05 1.786296e-05 1.916088e-05
```

```
[21] 1.973907e-05 2.139846e-05 2.368986e-05 2.582058e-05 2.586024e-05
```

```
[26] 2.972223e-05 4.067610e-05 4.086720e-05 4.282440e-05 4.522230e-05
```

# You can see these P values are different from the original ones.

# If you want to add another column to the table to list the P values corrected by Bonferroni method:

```
NT<-data.frame(MCC,PA)
```

```
View(NT)
```

# You can see the column name for these P values is “PA”. If you want to change the column name, you may do:

```
colnames(NT)<-  
c("Gene_name","Gf1","Gf2","Gf3","Gf4","Gf5","Gf6","Spf1","Spf2","Spf3","Spf4","  
Spf5","Spf6","GF_mean","SPF_mean","P_value_t-test","Bonferroni_P")
```

**View(NT)**

# Now you changed the name of that column to “Bonferroni\_P”.

# Use the same data, you may do a FDR correction as well:

```
PFDR<-p.adjust(PS,method='fdr',n=30)
```

**PFDR**

```
[1] 3.740985e-07 3.740985e-07 3.869960e-07 4.837170e-07 5.319600e-07  
[6] 5.319600e-07 6.476533e-07 6.476533e-07 6.476533e-07 8.145027e-07  
[11] 8.145027e-07 8.282460e-07 8.282460e-07 8.282460e-07 8.282460e-07  
[16] 9.399557e-07 9.399557e-07 9.399557e-07 9.399557e-07 9.399557e-07  
[21] 9.399557e-07 9.726573e-07 1.029994e-06 1.034410e-06 1.034410e-06  
[26] 1.143163e-06 1.459543e-06 1.459543e-06 1.476703e-06 1.507410e-06
```

# You can see these P values are different from the original ones or the ones corrected by Bonferroni method.

### **The answer to the challenge question:**

If you accidentally choose to do a paired t-test instead of using an unpaired t-test, you may get more false-positive results (Type I errors).

### Group Assignment 7:

1. Use R to solve the same problem from Midterm II:

You are interested in whether or not Germ-free mice have the same lifespan as the SPF mice.

You randomly picked 10 Germ-free mice and 10 SPF mice. You recorded how long these mice lived and tabulate the data in file “MouseLifeSpan.csv”.

What kind of statistical test would you like to perform? (0.5pt)

What is your null hypothesis? (0.5pt)

Set your alpha.

Find the p-value, if you decide to do a statistical test to compute a p-value. Show your R commands and results. (3.5pts)

What is your conclusion? (0.5pt)

2. Multiple comparison corrections

- 1) Use FDR correction to adjust the P values in file

“MouseMultipleComparisonsPCorrectionsAssignment.csv”. Show your R commands. (1.5pts)

- 2) Add a new column of to the table to list the adjusted P values. Show your R commands. (1pt)

- 3) Change the column name for the FDR corrected P values to “FDR\_P” Show your R commands. (1pt)

- 4) Write a CSV file to save your result. Name the CSV file as

“MultipleComparisonCorrectionFDRAssignment”. Show your R commands. (1.5pts)