

Final

Jiaying Liu

December 14, 2018

1(a)

The table below summarize data from a clinical trial to compare response to two treatments (a test drug and a placebo).

In this dataset, we have a 2 by 2 contingency table for response and nonresponse of placebo and test_drug. We want to compare response in two group, so we can actually test the response rate in two group

Assume the response in placebo is p_1 ;

Assume the response in Treatment is p_2

So the Null hypothesis is $H_0: p_1 = p_2$ $H_a: p_1 \neq p_2$

(b)

$$\text{test statistic: } T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}}$$

(c)

pvalue is 3.122698e-07

```
n1=64
n2=60
p1=16/64
p2=40/60
t=(p1-p2)/(sqrt(p1*(1-p1)/n1+p2*(1-p2)/n2))
pvalue=2*(1-pnorm(abs(t),mean=0,sd=1))
cbind(t,pvalue)

##           t           pvalue
## [1,] -5.115887 3.122698e-07
```

(d)

From above p value = 3.122698×10^{-7} which is greater smaller than 0.05, so we can reject the null hypothesis. So we can say the response in placebo and test drug is different.

2

(a)

In this question, we want to test whether the students receiving the improved curriculum would perform better than the standard curriculum in their math scores.

Assume the mean of old curriculum is μ_1

Assume the mean of new curriculum is μ_2

So, for the two side test the Null hypothesis is $H_0: \mu_1 = \mu_2$ $H_a: \mu_1 \neq \mu_2$

(b)

The test statistic when variance are different

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \approx t_{df}$$

$$\text{where } df = \frac{(n_1-1)(n_2-1)}{(n_1-1)(1-c)^2 + (n_2-1)c^2}$$

$$c = \frac{s_1^2}{s_1^2 + s_2^2}$$

The test statistic when variances are the same:

$$T = \frac{\bar{X} - \bar{Y}}{s_p \sqrt{1/n_1 + 1/n_2}} \approx t_{n_1+n_2-2}$$

$$\text{where } s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

(c)

Both result the p value are about 0.004

```
m_old=41.52
m_new=51.48
n_old=32
n_new=30
```

```

sd_old=17.15
sd_new=11.01
sp=((n_old-1)*(sd_old)^2+(n_new-1)*(sd_new)^2)/(n_old+n_new-2)
t=(m_new-m_old)/(sqrt(sp)*sqrt(1/n_old+1/n_new))
pvalue=1-pt(t,df=(n_old+n_new-2))
cbind(t,pvalue)

##           t           pvalue
## [1,] 2.700961 0.004486804

t=(m_new-m_old)/(sqrt((sd_old)^2/n_old+(sd_new)^2/n_new))
c=((sd_old)^2/n_old)/(((sd_old)^2/n_old)+((sd_new)^2/n_new))
df=((n_old-1)*(n_new-1))/((n_old-1)*(1-c)^2+(n_new-1)*c^2)
pvalue=1-pt(t,df)
cbind(t,pvalue)

##           t           pvalue
## [1,] 2.738083 0.004192562

```

(d)

From above p value = which is greater smaller than 0.05, so we can reject the null hypothesis. So, we can say the the performance of new curriculum is different from old curriculum. And if we want to test which one is better, we can see the t value is positive value, so the mean of new curriculum is better than old curriculum.

3

(a)

$$Y = UX$$

U is a binary outcome with $P(U = 1) = p, P(U = 0) = 1 - p$

$$\text{so } E(U) = p * 1 + (1 - p) * 0 = p$$

$$E(U^2) = p * 1 + (1 - p) * 0 = p$$

$$V(U) = p(1 - p)$$

Because ux is two independent variable so

$$E(Y) = E(UX) = E(U)E(X) = p\mu$$

$$Var(Y) = Var(UX) = E(U^2X^2) - (E(UX))^2$$

$$= var(U)var(X) + Var(U)(E(X))^2 + Var(X)(E(U))^2$$

$$= p(1 - p)\sigma^2 + p(1 - p)\mu^2 + p^2\sigma^2$$

$$= p\sigma^2 + p(1 - p)\mu^2$$

(b)

We already know X from $SC(p_1, \mu_1, \sigma_1)$ and Y from $SC(p_2, \mu_2, \sigma_2)$, $\delta = E(X) - E(Y)$. Suppose we wanted to test the null hypothesis $H_0: \delta = 0$ at a significance level of α with power $1 - \beta$ at $\delta = \delta_1 > 0$. So with the formula in the lecture notes, we can divide the sample as

$$n = \frac{Var(X) + Var(Y)}{\delta_1^2} (Z_{1-\beta} + Z_{1-\alpha/2})^2$$
$$= \frac{p_1\sigma_1^2 + p_1(1-p_1)\mu_1^2 + p_2\sigma_2^2 + p_2(1-p_2)\mu_2^2}{(p_1\mu_1 - p_2\mu_2)^2} (Z_{1-\beta} + Z_{1-\alpha/2})^2$$

(c)

```
p1=0.8
p2=0.85
lam1=1
lam2=1.1
```

```
m1=1/lam1
m2=1/lam2
```

```
s1=1/lam1^2
s2=1/lam2^2
```

```
v1=p1*s1+p1*(1-p1)*m1^2
v2=p2*s2+p2*(1-p2)*m2^2
c1=p1*m1-p2*m2
```

```
n=(v1+v2)/(c1^2)*((1.97+1.282)^2)
n
```

```
## [1] 24981.27
```

So the n is 24981 and the sample size is $2*24981=49962$

(d)

A type I error occurs if we reject the null hypothesis when the null hypothesis is true.

A power occurs if we reject the null hypothesis when the alternative hypothesis is true.

So we have the simulation showing below.

```

p1=0.8
p2=0.8
lam1=1
lam2=1
n1=10000 # testing 10,000 times
t1err=0
for (i in 1:n1){
  n=24981
  z1=rexp(n,rate=lam1)
  u1=rbinom(n,1,p1)
  x=z1*u1

  z2=rexp(n,rate=lam2)
  u2=rbinom(n,1,p2)
  y=z2*u2

  if (((t.test(x,y,mu=0,alternative = "two.sided",var.equal = FALSE))$
p.value)<=0.05) (t1err=t1err+1)
}
cat("Type I error rate in percentage is", (t1err/n1)*100,"%")

## Type I error rate in percentage is 4.81 %


p1=0.8
p2=0.85
lam1=1
lam2=1.1
m1=1/lam1
m2=1/lam2

n1=10000 # testing 10,000 times
t2err=0
for (i in 1:n1){
  n=24981
  z1=rexp(n,rate=lam1)
  u1=rbinom(n,1,p1)
  x=z1*u1

  z2=rexp(n,rate=lam2)
  u2=rbinom(n,1,p2)
  y=z2*u2

  if (((t.test(x,y, mu=0))$p.value)<=0.05) (t2err=t2err+1)
}
cat("power is", (t2err/n1)*100,"%")

## power is 89.86 %

```

4

task1

```
/*1.1*/  
proc means data=final MEAN STDERR;  
var HbA1c_level;  
class Visit_number Control_treatment;  
run;
```

The SAS System

The MEANS Procedure

Analysis Variable : HBA1C_level				
Visit_number	Control_treatment	N Obs	Mean	Std Error
1	Control	188	8.3290323	0.0826197
	Treatment	200	8.3939394	0.0837901
2	Control	186	8.6579235	0.1000459
	Treatment	192	8.7305263	0.0941217
3	Control	184	9.0774725	0.1132738
	Treatment	186	8.9432432	0.1014603
4	Control	180	9.1156425	0.1135777
	Treatment	180	9.0622222	0.1087088
5	Control	175	9.1620690	0.1222654
	Treatment	179	8.9268571	0.1085785
6	Control	164	9.0327160	0.1158673
	Treatment	166	8.8366460	0.1168847
7	Control	147	8.9618056	0.1262005
	Treatment	144	8.7929078	0.1082870
8	Control	110	8.8268519	0.1337585
	Treatment	103	8.6718447	0.1343981
9	Control	48	9.0375000	0.2263969
	Treatment	56	8.6875000	0.2022770

```

To test the different in mean
Two-sample t test
/*1.2*/
proc ttest data=final;
class Control_treatment;
var HbA1c_level;
run;

```

- Comparing different in control and treatment

The SAS System							
The TTEST Procedure							
Variable: HBA1C_level							
Control_treatment	N	Mean	Std Dev	Std Err	Minimum	Maximum	
Control	1366	8.8964	1.4713	0.0398	5.4000	16.9000	
Treatment	1389	8.7911	1.3782	0.0370	5.6000	16.8000	
Diff (1-2)		0.1053	1.4251	0.0543			

Control_treatment	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
Control		8.8964	8.8183	8.9745	1.4713	1.4181	1.5286
Treatment		8.7911	8.7186	8.8637	1.3782	1.3288	1.4315
Diff (1-2)	Pooled	0.1053	-0.00121	0.2117	1.4251	1.3884	1.4638
Diff (1-2)	Satterthwaite	0.1053	-0.00127	0.2118			

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	2753	1.94	0.0527
Satterthwaite	Unequal	2734.7	1.94	0.0528

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	1365	1388	1.14	0.0154

From the above table for the result of t sample t-test, we can see that from the third table, both pooled and Satterthwaite test show that the p-value is greater than 0.05 which mean the mean of HBA1C_level in treatment group and control group are the same. Also, we can see that from the fourth table, it shows the test for equality of variance, and the p-value is 0.154 smaller than 0.05, so we can say the variances in two groups are different, so we use the Satterthwaite test. And the result show there are same in two group.

- Comparing different in control and treatment in each visit

```
library(reshape)
final=read.csv("D:/GWU-A1C-Graphs.csv", header = TRUE)
final_trt=final[final[, "Control.treatment"]=="Treatment", ]
final_ctr=final[final[, "Control.treatment"]=="Control", ]

## t-test
t=c()
for (i in 1:9){
  t[i]=t.test(final_trt[which(final_trt$Visit.number==i),2],final_ctr[which(final_ctr$Visit.number==i),2])$p.value
}
t

## [1] 0.5815507 0.5974327 0.3779963 0.7342190 0.1512219 0.2344077 0.3106682
## [8] 0.4145858 0.2517682
```

Both p-value for visit in two group are greater than 0.05, so there are no different between each visit in each group.


```

/*1.3 Two sample Wilcoxon Rank test*/
proc npar1way data=final wilcoxon;
class Control_treatment;
var HbA1c_level;
run;

```

- Comparing different in control and treatment

The SAS System

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable HBA1C_level Classified by Variable Control_treatment					
Control_treatment	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Control	1366	1921362.50	1882348.0	20868.0661	1406.56113
Treatment	1389	1875027.50	1914042.0	20868.0661	1349.91181
Average scores were used for ties.					

Wilcoxon Two-Sample Test	
Statistic	1921362.5000
Normal Approximation	
Z	1.8696
One-Sided Pr > Z	0.0308
Two-Sided Pr > Z	0.0615
t Approximation	
One-Sided Pr > Z	0.0308
Two-Sided Pr > Z	0.0617
Z includes a continuity correction of 0.5.	

Kruskal-Wallis Test	
Chi-Square	3.4953
DF	1
Pr > Chi-Square	0.0615

From the above figure, it shows the result of Wilcoxon rank test. Based on the two-side test, we can see the p-value of normal approximation and t approximation are both equal to 0.06 which is greater than 0.05. Again, we cannot reject the null hypothesis which mean the mean of two treatment are same.

- Comparing different in control and treatment in each visit

```
## Wilcoxon rank sum test
w=c()
for (i in 1:9){
w[i]=wilcox.test(final_trt[which(final_trt$Visit.number==i),2],final_ctr[which(final_ctr$Visit.number==i),2])$p.value
}
w

## [1] 0.79841141 0.52683378 0.44754128 0.74241548 0.17020682 0.0900438
2
## [7] 0.48364829 0.27122339 0.23632318
```

Both p-value for visit in two group are greater than 0.05, so there are no different between each visit in each group.

```
/*1.4 Permutation test*/
proc multtest data=final permutation;
class Control treatment;
test mean(HbA1c_level);
run;
```

- Comparing different in control and treatment

The SAS System

The Multtest Procedure

Model Information	
Test for continuous variables	Mean t-test
Degrees of Freedom Method	Pooled
Tails for continuous tests	Two-tailed
Strata weights	None
P-value adjustment	Permutation
Center continuous variables	No
Number of resamples	20000
Seed	850811001

Adjusted Contrast Coefficients			
Contrast		Control_treatment	
		Control	Treatment
Trend	Centered	-0.5	0.5

Continuous Variable Tabulations				
Variable	Control_treatment	NumObs	Mean	Standard Deviation
HBA1C_level	Control	1366	8.8964	1.4713
HBA1C_level	Treatment	1389	8.7911	1.3782

p-Values			
Variable	Contrast	Raw	Permutation
HBA1C_level	Trend	0.0527	0.0527

Above figure show the result mean t test of permutation test. The p-value of permutation test is 0.0527 which is also greater than 0.05. So, under the permutation test, the result is also the same, there are not different of mean between two group.

- Comparing different in control and treatment in each visit

```
## permutation test
permu_test=function(x,y){
  nx=length(x)
  ny=length(y)
  n=nx+ny
  T0=t.test(x,y,var.equal = TRUE)$statistic
  xy=c(x,y)
  M=9999
  T=0
  for (i in 1:M){
    permxy=sample(xy,n)
    permx=permxy[c(1:nx)]
    permy=permxy[-c(1:nx)]
    T[i]=t.test(permx,permy,var.equal = TRUE)$statistic
  }
  pvalue=sum(abs(T)>=abs(T0))/M
  return(pvalue)
}
p=c()
for (i in 1:9){
  p[i]=permu_test(final_trt[which(final_trt$Visit.number==i),2],final_ctr
[which(final_ctr$Visit.number==i),2])
}
p

## [1] 0.5917592 0.5986599 0.3807381 0.7355736 0.1487149 0.2345235 0.31
38314
## [8] 0.4170417 0.2586259
```

So, with above three kind of test from parameter to non-parameter methods, the p value both greater than 0.05 and the result both show that there are no different of mean HbA1c-level between treatment and control group in each visit.

Task2

In this task, we are going to test the change in HbA1c level from baseline, so the first thing we need to do is to calculate the different of each visit with the baseline.

```
/*2*/  
proc sql;  
create table final2 as  
select PATID,Visit_number,HBA1C_level as baseline  
from final  
where Visit_number=1  
group by PATID;
```

```
proc sql;  
create table final3 as  
select a.*,b.baseline  
FROM final as a, final2 as b  
WHERE a.PATID=b.PATID;
```

```
data final3;  
set final3;  
different=HBA1C_level-baseline;  
run;
```

```
data final3;  
set final3;  
if Visit_number=1 then delete;  
run;
```

	Visit_number	HBA1C_level	PATID	Control_treatment	Age	baseline	different
1	2	8.8	1800	Control	11.1266256	8.9	-0.1
2	3	9.6	1800	Control	11.1266256	8.9	0.7
3	4	10.5	1800	Control	11.1266256	8.9	1.6
4	5	10.2	1800	Control	11.1266256	8.9	1.3
5	6	9.6	1800	Control	11.1266256	8.9	0.7
6	7	8.9	1800	Control	11.1266256	8.9	0
7	2	11.5	1801	Control	11.56468173	10.9	0.6
8	3	11.6	1801	Control	11.56468173	10.9	0.7
9	4	11.6	1801	Control	11.56468173	10.9	0.7
10	5	11.5	1801	Control	11.56468173	10.9	0.6
11	6	9.7	1801	Control	11.56468173	10.9	-1.2
12	7	9.8	1801	Control	11.56468173	10.9	-1.1
13	2	7.1	1802	Treatment	12.32854209	6.9	0.2
14	3	7.1	1802	Treatment	12.32854209	6.9	0.2
15	4	8.3	1802	Treatment	12.32854209	6.9	1.4
16	5	7.5	1802	Treatment	12.32854209	6.9	0.6
17	6	7.3	1802	Treatment	12.32854209	6.9	0.4
18	7	6.8	1802	Treatment	12.32854209	6.9	-0.1
19	8	7.3	1802	Treatment	12.32854209	6.9	0.4
20	2	9	1803	Control	13.21013005	7.7	1.3
21	3	7.7	1803	Control	13.21013005	7.7	0
22	4	8.1	1803	Control	13.21013005	7.7	0.4
23	5	8.5	1803	Control	13.21013005	7.7	0.8
24	6	8.6	1803	Control	13.21013005	7.7	0.9
25	7	9	1803	Control	13.21013005	7.7	1.3
26	8	8.8	1803	Control	13.21013005	7.7	1.1

With above statement, I was able to create the different of HBA1C between baseline and follow up visit in each PATID.

```

/*2.1 Mean and SE*/
proc means data=final3 MEAN STDERR;
var different;
class Visit_number Control_treatment;
run;

```

The SAS System

The MEANS Procedure

Analysis Variable : different				
Visit_number	Control_treatment	N Obs	Mean	Std Error
2	Control	186	0.3237569	0.0844198
	Treatment	192	0.3920213	0.0653892
3	Control	184	0.7155556	0.1027457
	Treatment	186	0.5890710	0.0836115
4	Control	180	0.7644068	0.1045120
	Treatment	180	0.6764045	0.0872099
5	Control	175	0.8578035	0.1166623
	Treatment	179	0.5693642	0.0919695
6	Control	164	0.7279503	0.1213255
	Treatment	166	0.5157233	0.0954149
7	Control	147	0.6741259	0.1230450
	Treatment	144	0.4892086	0.0922934
8	Control	110	0.6757009	0.1310056
	Treatment	103	0.3900990	0.1350944
9	Control	48	1.0127660	0.2249189
	Treatment	56	0.3870370	0.1899289

```
/*2.2 Two sample t test*/
proc ttest data=final3;
class Control_treatment;
var different;
run;
```

- Comparing different in control and treatment

The SAS System						
The TTEST Procedure						
Variable: different						
Control_treatment	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	1169	0.6883	1.4165	0.0414	-3.7000	10.0000
Treatment	1175	0.5197	1.1548	0.0337	-3.2000	6.1000
Diff (1-2)		0.1685	1.2920	0.0534		

Control_treatment	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
Control		0.6883	0.6070 0.7696	1.4165	1.3613 1.4764
Treatment		0.5197	0.4536 0.5858	1.1548	1.1099 1.2035
Diff (1-2)	Pooled	0.1685	0.0639 0.2732	1.2920	1.2560 1.3301
Diff (1-2)	Satterthwaite	0.1685	0.0638 0.2733		

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	2342	3.16	0.0016
Satterthwaite	Unequal	2246.2	3.16	0.0016

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	1168	1174	1.50	<.0001

Form the above result of two sample t test from the different between treatment and control group. With the test of equality of variances, we know that the variance of two group are different, so we use the outcome of Satterthwaite test. And we can see the p value of Satterthwaite is 0.0016 smaller than the 0.05, so we can reject the null hypothesis which meant there are a significant different of change for baseline and follow up in the treatment and control group.

- Comparing different in control and treatment in each visit

```
MyData=read.csv("d:/final3.csv")
MyData_trt=MyData[MyData[, "Control_treatment"]=="Treatment", ]
MyData_ctr=MyData[MyData[, "Control_treatment"]=="Control", ]

## t-test
t=c()
for (i in 2:9){
```

```
t[i]=t.test(MyData_trt[which(MyData_trt$Visit_number==i),7],MyData_ctr[
which(MyData_ctr$Visit_number==i),7])$p.value
}
t

## [1]          0.52306462 0.34032795 0.51838249 0.05304167 0.17015844
## [7] 0.23036235 0.13063401 0.03617129
```

However, when it comes to the mean of change in each visit of two group, there are some different. For the visit 2 to 8, there are no different because the p value greater than 0.05. But for visit 9, since the p value is 0.036 smaller thnn 0.05, there exist significant different between two group in the visit 9 compare with baseline.

```
/*2.3 Willcoxon*/
proc npar1way data=final3 wilcoxon;
class Control_treatment;
var different;
run;
```

- Comparing different in control and treatment

The SAS System					
The NPAR1WAY Procedure					
Wilcoxon Scores (Rank Sums) for Variable different Classified by Variable Control_treatment					
Control_treatment	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Control	1169	1423026.50	1370652.50	16380.5807	1217.30240
Treatment	1175	1325313.50	1377687.50	16380.5807	1127.92638
Average scores were used for ties.					

Wilcoxon Two-Sample Test	
Statistic	1423026.5000
Normal Approximation	
Z	3.1973
One-Sided Pr > Z	0.0007
Two-Sided Pr > Z	0.0014
t Approximation	
One-Sided Pr > Z	0.0007
Two-Sided Pr > Z	0.0014
Z includes a continuity correction of 0.5.	

Kruskal-Wallis Test	
Chi-Square	10.2229
DF	1
Pr > Chi-Square	0.0014

From the above result of Wilcoxon score rand test, based on the two-sided test, we can see both p value of normal approximation and t approximation are equal to 0.0014 which is smaller than 0.05. So, again, with the Wilcoxon test, it shows that there exist the different of change in two group.

- Comparing different in control and treatment in each visit

```
w=c()
for (i in 2:9){
w[i]=wilcox.test(MyData_trt[which(MyData_trt$Visit_number==i),7],MyData_
_ctr[which(MyData_ctr$Visit_number==i),7])$p.value
}
w

## [1]      0.79458171 0.49480739 0.32325318 0.05709312 0.08856235
## [7] 0.37311802 0.06537382 0.03802862
```

From the Wilcoxon result, it also have the same result like t test, the p value of visit 2 to 8 is greater than 0.05 but the p value of visit 9 is smaller than 0.05, so visit 9 is only one that have the different of change in each level in two group.


```
/*2.4 Permutation test*/
proc multtest data=final3 permutation;
class Control_treatment;
test mean(different);
run;
```

- Comparing different in control and treatment

The SAS System

The Multtest Procedure

Model Information	
Test for continuous variables	Mean t-test
Degrees of Freedom Method	Pooled
Tails for continuous tests	Two-tailed
Strata weights	None
P-value adjustment	Permutation
Center continuous variables	No
Number of resamples	20000
Seed	989115001

Adjusted Contrast Coefficients			
Contrast		Control_treatment	
		Control	Treatment
Trend	Centered	-0.5	0.5

Continuous Variable Tabulations				
Variable	Control_treatment	NumObs	Mean	Standard Deviation
different	Control	1169	0.6883	1.4165
different	Treatment	1175	0.5197	1.1548

p-Values			
Variable	Contrast	Raw	Permutation
different	Trend	0.0016	0.0016

Above figure show the result mean t test of permutation test. The p-value of permutation test is 0.0016 which is also smaller than 0.05. So, under the permutation test, the result is also the same, there exist different of change between two group.

- Comparing different in control and treatment in each visit

```
## permutation test
permu_test=function(x,y){
nx=length(x)
```

```

ny=length(y)
n=nx+ny
T0=t.test(x,y,var.equal = TRUE)$statistic
xy=c(x,y)
M=9999
T=0
for (i in 1:M){
  permxy=sample(xy,n)
  permx=permxy[c(1:nx)]
  permy=permxy[-c(1:nx)]
  T[i]=t.test(permx,permy,var.equal = TRUE)$statistic
}
pvalue=sum(abs(T)>=abs(T0))/M
return(pvalue)
}
p=c()
for (i in 2:9){
  p[i]=permu_test(MyData_trt[which(MyData_trt$Visit_number==i),7],MyData_
ctr[which(MyData_ctr$Visit_number==i),7])
}
p

## [1]          0.52025203 0.34543454 0.51275128 0.05190519 0.17421742
## [7] 0.23412341 0.13211321 0.03460346

```

Also, the result is the same like previous two test for the each visit in two group.

So, with above three kind of test from parameter to non-parameter methods, the p value of visit 2 to 8 both greater than 0.05 and only visit 9 show the different for its p value smaller than 0.05. So, there is only visit 9 have the significant different for the diff of baseline in treatment and control.

Task3

In this take, we are going to use a linear mixed effect model to test whether the two arms have the same rate of change.

For the mixed linear model:

let $Y = \text{HbA1c}$ $T = \text{visit}$ x : treatment=1; control= 0

So, we have

$$E(Y) = \beta_0 + \beta_1 X + \beta_2 T + \beta X * T$$

when $X=0$ control group $E(Y) = \beta_0 + \beta_2 T$

When $X=1$ treatment group $E(Y) = (\beta_0 + \beta_1) + (\beta_2 + \beta)T$

So, in order to test the effect of rate of change in treatment group and control group, we are going to test

$$H_0: \beta = 0$$

We can use the R to compute the mixed model, We set the random effect is visit.number and PATID

```
library(lme4)

## Loading required package: Matrix

m2=lmer(HBA1C.level~Control.treatment+Visit.number+Control.treatment*Visit.number+(Visit.number | PATID),data=MyData)
summary(m2)

## Linear mixed model fit by REML ['lmerMod']
## Formula:
## HBA1C.level ~ Control.treatment + Visit.number + Control.treatment *
##
## Visit.number + (Visit.number | PATID)
## Data: MyData
##
## REML criterion at convergence: 7928.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.5345 -0.4771 -0.0494  0.4109  5.8248
##
## Random effects:
##   Groups      Name                Variance Std.Dev. Corr
##   PATID      (Intercept)  1.16988   1.0816
##              Visit.number  0.02398   0.1549  -0.20
## Residual                        0.64166   0.8010
## Number of obs: 2755, groups: PATID, 388
##
## Fixed effects:
##                                     Estimate Std. Error t value
## (Intercept)                        8.51649    0.09218  92.386
## Control.treatmentTreatment          0.11409    0.12868   0.887
## Visit.number                       0.09619    0.01550   6.205
## Control.treatmentTreatment:Visit.number -0.04174    0.02188  -1.908
##
## Correlation of Fixed Effects:
##              (Intr) Cntr.T Vst.nm
## Cntrl.trtmT -0.716
## Visit.numbr -0.432  0.310
## Cntrl.tT:V.  0.306 -0.432 -0.709
```

From above information of estimation, we need to evaluate the significance of which is the effect of control.treatment:visit.number. we find out that the t value is -1.908

and the absolute value of t is 1.908 which is smaller than 1.96 (the value of $z_{1-\alpha/2}$) so we can not reject the original hypothesis which mean the two arms have same rate of change.