HW1

# 임상시험자료분석 II

182STG27

임지연

**1.** SCORE가 20보다 작으면 treatment의 효과가 있다고 알려져 있다고 가정 하자. 이 때 treatment A가 효과가 있는지를 알고자 한다.

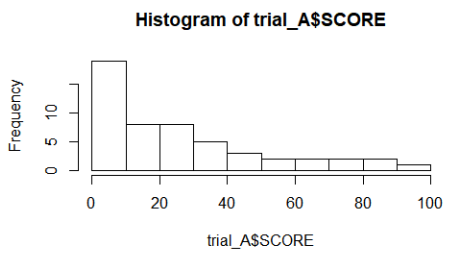
① 귀무가설과 대립가설을 쓰시오.
② 위의 가설을 그림으로 확인하고자 한다. 해당하는 그림을 그리시오
③ ① 의 가설 검정을 위하여 어떠한 방법을 써야 하는가? 기술하시오.
④ 위에서 기술한 방법을 SAS와 R을 각각 이용하여 결과를 얻은 후 이 두 결과를 비교하여 결론을 내리시오.

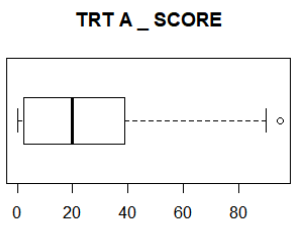
## R

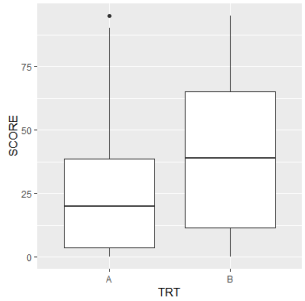
### 1. CODE

```
trial <- read.csv("C:/Users/jeeyeon/Desktop/data/ex3-1.csv")
trial_A <- trial %>% filter(TRT == "A"); trial_A
trial_B <- trial %>% filter(TRT == "B"); trial_B
ggplot(data = trial_A, aes(SCORE))+geom_histogram()   #histogram
boxplot(trial_A$SCORE,horizontal = TRUE) #boxplot
ggplot(aes(y = SCORE, x = TRT), data = trial) + geom_boxplot() #TRT 두집단비교
t.test(trial_A$SCORE,mu = 20 ,alternative ="less") #T.test
```

### 2. PLOT




### 3. RESULT OF T-TEST

```
> t.test(trial_A$SCORE,mu = 20 ,alternative ="less")

        One Sample t-test

data:  trial_A$SCORE
t = 1.7855, df = 51, p-value = 0.9599
alternative hypothesis: true mean is less than 20
95 percent confidence interval:
     -Inf 32.93428
sample estimates:
mean of x
 26.67308
```

t-test 결과 p-value = 0.9599로, 귀무가설을 기각하지 못한다. 따라서 TRT = A 그룹의 SCORE 점수가 20이상이라는 귀무가설을 채택하게 된다.

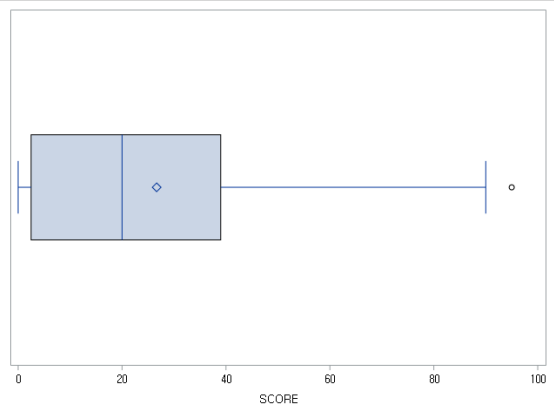
## SAS

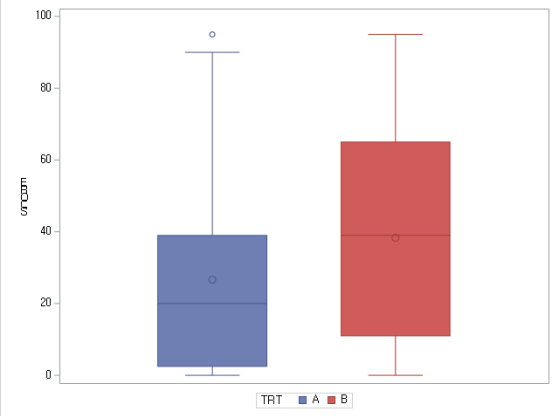
### 1. CODE

```
proc import
datafile="C:₩Users₩jeeyeon₩Desktop₩data₩ex3-1.csv"
out=trial;
getnames=yes;
run;

data trial_A;
set trial;
if TRT="A";run;

/* TRT_A boxplot   */
proc sgplot data=trial_A;
            hbox SCORE; run;
```

```
/* TRT boxplot   */
proc sgplot data=trial;
            vbox SCORE / group=TRT; run;

/*  T-test   */
ods graphics on;
   proc ttest h0=20 plots(showh0) sides=l data=trial_A;
      var SCORE;
   run;
ods graphics off;
```

## 2. PLOT




## 3. RESULT OF T-TEST

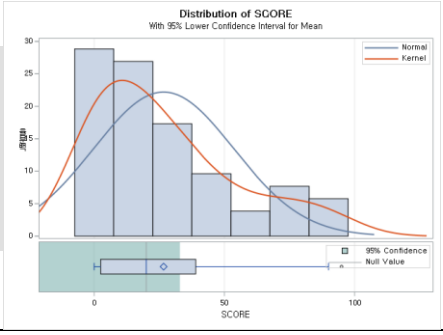
*The TTEST Procedure*

*Variable: SCORE*

| N | Mean | Std Dev | Std Err | Minimum | Maximum |
|---|---|---|---|---|---|
| 52 | 26.6731 | 26.9507 | 3.7374 | 0 | 95.0000 |

| Mean | 95% CL Mean | | Std Dev | 95% CL Std Dev | |
|---|---|---|---|---|---|
| 26.6731 | -Infty | 32.9343 | 26.9507 | 22.5860 | 33.4223 |

| DF | t Value | Pr < t |
|---|---|---|
| 51 | 1.79 | 0.9599 |




t-test 결과 p-value = 0.9599로, 귀무가설을 기각하지 못한다. 따라서 TRT = A 그룹의 SCORE 점수가 20이상이라는 귀무가설을 채택하게 된다.

## 결과 해석

① 문제에서 TRT_A 에 대해 SCORE 값의 평균이 20 이하이면 효과가 있다고 했으므로 이를 위한 가설은 귀무가설 (Ho) : mu ≥ 20 , 대립가설 (H1) : mu < 20 로 설정된다.

② SCORE 변수의 대략적인 형태를 보기 위하여 히스토그램과 Boxplot 을 그려서 시각화하였다. Boxplot 형태를 봤을 때, 두 집단의 차이가 어느 정도 있을 것이라고 판단하였다.

③ 위 가설을 검정하기 위하여 SAS, R 을 이용하여 one-sample t-test 한 후 결과를 비교하였다.

④ 결과는 p-value = 0.9599 로, 귀무가설을 기각하지 못한다. 따라서 TRT = A 그룹의 SCORE 점수가 20 이상이라는 귀무가설을 채택하게 된다. 따라서 Treatment A 의 효과는 없을 것이라고 결론지을 수 있다.

**2.** SCORE 변수를 이용하여

① treatment group 간에 차이가 있는지를 one-way ANOVA를 이용하여 분석하시오.

② treatment group과 center 간에 교호작용이 있는지를 알고자 한다. ANOVA를 이용하여 결론을 내리시오.

③ treatment group과 center간에 교호작용이 없다고 가정하고 treatment 와 center 간 차이를 알고자 한다. ANOVA를 이용하여 결론을 내리시오.

④ ③에서 treatment 또는 center 간 차이가 있다는 결론이 나왔을 경우 어떻게 해야 하는가?

**R**

```
trial <- trial %>% mutate(center_c = as.factor(CENTER) )
t.test(trial$SCORE[trial$TRT == 'A'],trial$SCORE[trial$TRT == 'B'])
anova(lm( SCORE ~ TRT, data = trial)) #alpha = 0.05 -> Ho reject (Ho accept)
anova(lm( SCORE ~ TRT * center_c, data = trial)) #alpha = 0.05 -> Ho do not reject (H1 accept)
anova(lm( SCORE ~ TRT + center_c, data = trial))   #alpha = 0.05 -> Trt : Ho reject(H1 accept)
                                                   #alpha = 0.05 -> Center :Ho do not reject(Ho accpet)
```

```
> anova(lm( SCORE ~ TRT, data = trial))
Analysis of Variance Table

Response: SCORE
          Df Sum Sq Mean Sq F value Pr(>F)
TRT        1   3394  3393.6  4.2375 0.0422 *
Residuals 98  78484   800.9
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

One-way ANOVA 결과, P-value = 0.0422로, 유의수준 0.05에서 trt의 효과가 있다는 대립가설을 채택하여 효과가 있다고 할 수 있다.

```
> anova(lm( SCORE ~ TRT*center_c, data = trial))
Analysis of Variance Table

Response: SCORE
             Df Sum Sq Mean Sq F value Pr(>F)
TRT           1   3394  3393.6  4.2047 0.0431 *
center_c      2   1145   572.5  0.7093 0.4946
TRT:center_c  2   1472   736.1  0.9121 0.4052
Residuals    94  75867   807.1
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

One-way ANOVA 결과, Treatment와 Center의 교호효과의 P-value = 0.4052로, 유의수준 0.05에서 교호효과가 없다는 귀무가설을 채택한다.
따라서 Treatment와 Center의 교호효과는 유의하지 않다.

```
> anova(lm( SCORE ~ TRT + center_c, data = trial))
Analysis of Variance Table

Response: SCORE
          Df Sum Sq Mean Sq F value  Pr(>F)
TRT        1   3394  3393.6  4.2124 0.04285 *
center_c   2   1145   572.5  0.7106 0.49389
Residuals 96  77339   805.6
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

One-way ANOVA 결과, Treatment의 P-value =0.042로 유의수준 0.05에서 귀무가설을 기각하여 Treatment의 효과가 있다고 할 수 있다.
Center의 P-value =0.494로 유의수준 0.05에서 귀무가설을 기각하지 못하여, Center의 효과가 없다는 귀무가설을 채택한다.

**SAS**

```
/*  one-way ANOVA  */
proc glm data=trial;
        class TRT CENTER;
        model SCORE = TRT / SS3; run;
/*  교호효과검정  */
proc glm data=trial;
        class TRT CENTER;
        model SCORE=TRT CENTER TRT*CENTER; run;
/*  TRT, CENTER효과검정  */
proc glm data=trial;
        class TRT CENTER;
        model SCORE=TRT CENTER; run;
```

*The GLM Procedure*

| Class Level Information | |
|---|---|
| Class | Levels Values |
| TRT | 2 A B |

| | |
|---|---|
| Number of Observations Read | 100 |
| Number of Observations Used | 100 |

페이지 나누기

*The GLM Procedure*

*Dependent Variable: SCORE*

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 1 | 3393.60103 | 3393.60103 | 4.24 | 0.0422 |
| Error | 98 | 78484.10897 | 800.85825 | | |
| Corrected Total | 99 | 81877.71000 | | | |

| R-Square | Coeff Var | Root MSE | SCORE Mean |
|---|---|---|---|
| 0.041447 | 87.69581 | 28.29944 | 32.27000 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| TRT | 1 | 3393.601026 | 3393.601026 | 4.24 | 0.0422 |

One-way ANOVA 결과, P-value = 0.0422로, 유의수준 0.05에서 trt의 효과가 있다는 대립가설을 채택하여 효과가 있다고 할 수 있다.

*The GLM Procedure*

*Dependent Variable: SCORE*

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 5 | 6010.89531 | 1202.17906 | 1.49 | 0.2006 |
| Error | 94 | 75866.81469 | 807.09377 | | |
| Corrected Total | 99 | 81877.71000 | | | |

| R-Square | Coeff Var | Root MSE | SCORE Mean |
|---|---|---|---|
| 0.073413 | 88.03655 | 28.40940 | 32.27000 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| TRT | 1 | 3393.601026 | 3393.601026 | 4.20 | 0.0431 |
| CENTER | 2 | 1145.014504 | 572.507252 | 0.71 | 0.4946 |
| TRT*CENTER | 2 | 1472.279781 | 736.139891 | 0.91 | 0.4052 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| TRT | 1 | 2706.907813 | 2706.907813 | 3.35 | 0.0702 |
| CENTER | 2 | 999.961195 | 499.980597 | 0.62 | 0.5404 |
| TRT*CENTER | 2 | 1472.279781 | 736.139891 | 0.91 | 0.4052 |

One-way ANOVA 결과, Treatment와 Center의 교호효과의 P-value = 0.4052로, 유의수준 0.05에서 교호효과가 없다는 귀무가설을 채택한다.
따라서 Treatment와 Center의 교호효과는 유의하지 않다.

*The GLM Procedure*

*Dependent Variable: SCORE*

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 3 | 4538.81553 | 1512.87184 | 1.88 | 0.1385 |
| Error | 96 | 77339.09447 | 805.61557 | | |
| Corrected Total | 99 | 81877.71000 | | | |

| R-Square | Coeff Var | Root MSE | SCORE Mean |
|---|---|---|---|
| 0.055432 | 87.95590 | 28.38337 | 32.27000 |

| Source | DF | Type I SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| TRT | 1 | 3393.601026 | 3393.601026 | 4.21 | 0.0429 |
| CENTER | 2 | 1145.014504 | 572.507252 | 0.71 | 0.4939 |

| Source | DF | Type III SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| TRT | 1 | 3572.926244 | 3572.926244 | 4.44 | 0.0378 |
| CENTER | 2 | 1145.014504 | 572.507252 | 0.71 | 0.4939 |

One-way ANOVA 결과, Treatment의 P-value =0.0429로 유의수준 0.05에서 귀무가설을 기각하여 Treatment의 효과가 있다고 할 수 있다. Center의 P-value =0. 494로, 유의수준 0.05에서 귀무가설을 기각하지 못하여, Center의 효과가 없다는 귀무가설을 채택한다.

## 결과 해석

① treatment group 간에 차이가 있는지를 one-way ANOVA 를 이용하여 분석한 결과, 유의수준 0.05 에서 trt 의 효과가 있다는 대립가설을 채택하여 treatment group 간에 차이가 있다고 할 수 있다.

② treatment group과 center 간에 교호효과에 대한 P-value = 0.187로, 유의수준 0.05에서 교호효과가 없다.

③ treatment group과 center간에 교호작용이 없다고 가정하고 treatment 와 center 간 차이를 알기 위해 ANOVA를 이용하여 분석한 결과, Treatment의 효과가 있고, Center의 효과는 없다는 귀무가설을 채택한다.

④ 만약, treatment 효과가 있다는 결론이 나왔을 경우 방안은 다음과 같다.

| TRT | CENTER | 목적 : treatment A 가 효과가 있는가 ? |
|---|---|---|
| O | O | treatment A 가 효과가 있는지를 알기 위하여 Center 의 효과를 제어한 후 Treatment 의 효과를 다시 살펴본다. |
| O | X | center 간 차이가 없다는 것은 바람직한 결론으로, Treatment A 의 효과를 T.test 를 통해 살펴본다. |
| X | O | center 간의 차이가 있다는 것은 편향이 발생했다는 의미이기 때문에 이를 해결할 수 있는 구체적인 방안이 필요하다. Center 간의 차이가 없는 것이 바람직한 결론이므로 원인이 무엇인지 구체적으로 탐색한 후, Center 의 효과가 없게 다시 실험설계한 후 임상시험을 진행한다. |