

## 실험 계획 및 실습

### Homework #1) 2014-16757 김보창

## 1 Q1

2.15

(a) : output을 보면 Two sample T-test의 P-value가 0.001로 나와있는데, P-value의 의미는  $H_0$ 를 기각하지 않기 위한 최소의 신뢰수준  $\alpha$ 의 값을 의미하므로, 0.05 level에서  $\alpha=0.05$ 이므로, P-value가 0.05보다 작으므로  $H_0$ 를 기각한다.

(b) : 결과에서, T-test of difference = 0 (vs not =)의 부분을 보아,  $H_0 : u_1=u_2$ ,  $H_1 : u_1 \neq u_2$ 에 대해 test를 진행하는 것임을 알 수 있으므로, 이는 two-sided test임을 알 수 있다.

(c) : output을 보면  $y_1, y_2$ 가 normal distribution을 따른다 가정하고 등분산 가정을 하고 test를 진행하고 있음을 알 수 있고, 따라서 output에서의 test statistic은

$$T = \frac{\bar{y}_1 - \bar{y}_2 - (u_1 - u_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

와 같이 됨을 알 수 있고, output에서  $n_1 = 20, n_2 = 20, S_p = 2.1277$ 이 됨을 안다.  
이때,  $H_0$ 하에서

$$T_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(38)$$

임을 알고, 이때의  $T_0 = -3.47$ 임을 안다.

이제,  $H_0 : u_1 - u_2 - 2 = 0, H_1 : u_1 - u_2 - 2 \neq 0$ 에 대해 test statistic을 세우면,  
 $E(\bar{y}_1 - \bar{y}_2 - 2) = u_1 - u_2 - 2$ 에서,  $\bar{y}_1 - \bar{y}_2 - 2$ 는  $u_1 - u_2 - 2$ 의 unbiased estimator임을 안다.  
또한,

$$T = \frac{\bar{y}_1 - \bar{y}_2 - 2 - (u_1 - u_2 - 2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

임을 알고 있고,  $H_0$ 가 사실일때

$$T'_0 = \frac{\bar{y}_1 - \bar{y}_2 - 2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(38)$$

임을 아므로,

이 값을 구해서 two-side t test를 진행하면 될것이다. 즉,  $|T'_0| > t_{\frac{\alpha}{2}}(38)$ 임을 체크하면 된다.

따라서  $T'_0 = -3.47 - \frac{2}{2.1277\sqrt{\frac{1}{10}}} = -6.44$ 로, 이 값의 절대값이  $\alpha = 0.05$ 일때  $t_{0.025}(38) = 2.02$ 보다 크므로,  $H_0$ 를 기각하게 된다.

(d) : (c)와 같지만  $H_1 : u_1 - u_2 - 2 < 0$ 으로  $H_1$ 이 바뀐것이 다르다. 귀무가설이 같으므로, 앞의 통계량을 그대로 사용할 수 있다.

따라서 앞의 통계량을 그대로 사용하되, one-side t test를 진행하면 될것이다. 즉,  $T'_0 < t_{1-\alpha}(38) = -t_\alpha(38)$ 임을 체크하면 된다.

이때,  $\alpha = 0.05$ 일때  $t_{0.05}(38) = 1.69$ 로,  $-6.44 < -1.69$ 에서 역시  $H_0$ 를 기각할 수 있다.

(e) :  $H_0 : u_1 - u_2 \leq 0, H_1 : u_1 - u_2 > 0$ 일때, 95% upper confidence bound를 구하자.

## 실험 계획 및 실습

### Homework #1) 2014-16757 김보창

결국,  $u_1 - u_2$ 가 특정값보다 작아지는 신뢰도 95% 구간을 구하고 싶기 때문에,  $P(u_1 - u_2 \leq U) = 0.95$  인  $U$ 를 구하자.

$$T = \frac{\bar{y}_1 - \bar{y}_2 - (u_1 - u_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

임을 아므로,  $P(\frac{\bar{y}_1 - \bar{y}_2 - (u_1 - u_2)}{S_p \sqrt{\frac{1}{10}}} \geq t_{0.95}(38)) = 0.95$ 에서,

$P(\bar{y}_1 - \bar{y}_2 - S_p \sqrt{\frac{1}{10}} t_{0.95}(38) \geq u_1 - u_2) = 0.95$  이므로,

$$U = \bar{y}_1 - \bar{y}_2 - S_p \sqrt{\frac{1}{10}} t_{0.95}(38) = -2.33341 - \frac{2.1277}{\sqrt{10}}(-1.686) = -1.199$$

임을 알 수 있다.

(f) : P-value의 정의에 의해,  $H_0 : u_1 - u_2 - 2 = 0, H_1 : u_1 - u_2 - 2 \neq 0$  일때 이를 구하면, (c)에서와 같이,  $T'_0 = -3.47 - \frac{2}{2.1277\sqrt{\frac{1}{10}}} = -6.44$  이므로,  $P(|T| > t_{\frac{\alpha}{2}}(38)) = P(|T| > 6.44)$  가 되는 알파가 곧 p-value가 된다. 이를 직접 찾기에는 너무 힘들므로, 프로그램을 이용할것이다.

R의 BSDA 패키지를 이용하였다. `install.packages('BSDA')` 문으로 설치할 수 있다.

```
> library(BSDA)
필요한 패키지를 로딩중입니다: lattice

다음의 패키지를 부속합니다: 'BSDA'

The following object is masked from 'package:datasets':

  orange

warning message:
패키지 'BSDA'는 R 버전 3.6.3에서 작성되었습니다
> tsum.test(mean.x = 50.19, mean.y = 52.52, s.x = 1.71, s.y = 2.48, n.x =
  20, n.y = 20, mu = 2)

welch Modified Two-sample t-Test

data: summarized x and y
t = -6.4282, df = 33.736, p-value =
2.492e-07
alternative hypothesis: true difference in means is not equal to 2
95 percent confidence interval:
-3.6992971 -0.9607029
sample estimates:
mean of x mean of y
  50.19      52.52

> |
```

$u = 2$ 로 주어 가설에 맞게 검정하였다.

이때의 p-value는  $2.492 \times 10^{-7}$ 로, 굉장히 작음을 알 수 있다.

## 2 Q2

2.34

## 실험 계획 및 실습

### Homework #1) 2014-16757 김보창

$H_0 : u_1 - u_2 = 0, H_1 : u_1 - u_2 \neq 0$  에서,

각 girder간에 method에 따른 effect가 다르게 계산되므로, 다음과 같이 모델링 할 수 있다.

$$y_{ij} = u_i + \beta_j + \epsilon_{ij} \quad i = 1, 2 \quad j = 1, 2, \dots, 9$$

이경우, paired t-test를 하는것이 적절하고,

$d_j = y_{1j} - y_{2j} \quad j = 1, 2, \dots, 9$  에서,  $E(d_j) = u_1 - u_2$ 가 성립하고,  $d_j \stackrel{\text{i.i.d.}}{\sim} N(u_1 - u_2, \sigma^2)$  로 서로 독립이고, 정규분포를 따른다고 가정하자.

$$T = \frac{\bar{d} - (u_1 - u_2)}{\frac{S_d}{\sqrt{n}}} \sim t(n-1)$$
$$S_d^2 = \frac{\sum_{k=1}^n (d_k - \bar{d})^2}{n-1}$$

임을 아므로,

이를 이용하여 가설검정을 진행할 것이다.

(a) :  $\alpha = 0.05$ 일때, 가설검정을 진행하자.

$H_0$ 가 사실일때,  $T_0 = \frac{\bar{d}}{\frac{S_d}{\sqrt{n}}} \sim t(n-1)$  가 성립하므로, 양측검정에서

$|T_0| > t_{0.025}(8)$  인지를 체크하면 된다. 이때의  $T_0$  를 구하면,

```
> a <- c(1.186, 1.151, 1.322, 1.339, 1.200, 1.402, 1.365, 1.537, 1.559)
> b <- c(1.061, 0.992, 1.063, 1.062, 1.065, 1.178, 1.037, 1.086, 1.052)
> d <- a-b
> mean(d)
[1] 0.2738889
> sd(d)
[1] 0.1350994
```

에서,  $T_0 = 0.2739 / (0.1351) * \sqrt{9} = 6.082$  가 되고, 이는  $t_{0.025}(8) = 2.306$ 보다 크게되어 귀무가설  $H_0$ 를 기각할 수 있다.

(b) : P-value를 구하기 위해, 1-(f)에서 처럼 프로그램을 이용할 것이다. R의 내장 라이브러리를 이용하였다.

```
> t.test(a,b,paired = TRUE)

Paired t-test

data: a and b
t = 6.0819, df = 8, p-value = 0.0002953
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.1700423 0.3777355
sample estimates:
mean of the differences
      0.2738889
```

P-value가 0.0002953인것을 확인할 수 있다.

(c) : 위  $H_0$ 과  $H_1$ 를 사용하여 차이의 평균의 95% 신뢰구간을 구하자. 즉,  $u_1 - u_2$ 의 신뢰구간을 구하자.

$$T = \frac{\bar{d} - (u_1 - u_2)}{\frac{S_d}{\sqrt{n}}} \sim t(n-1)$$

임을 아므로,

## 실험 계획 및 실습

### Homework #1) 2014-16757 김보창

이의 신뢰구간은  $P(|\frac{\bar{d} - (u_1 - u_2)}{\frac{S_d}{\sqrt{n}}}| \leq t_{\frac{\alpha}{2}}(n-1)) = \alpha$  에서,

$$\bar{d} - t_{0.025}(8) \frac{S_d}{\sqrt{9}} \leq u_1 - u_2 \leq \bar{d} + t_{0.025}(8) \frac{S_d}{\sqrt{9}}$$

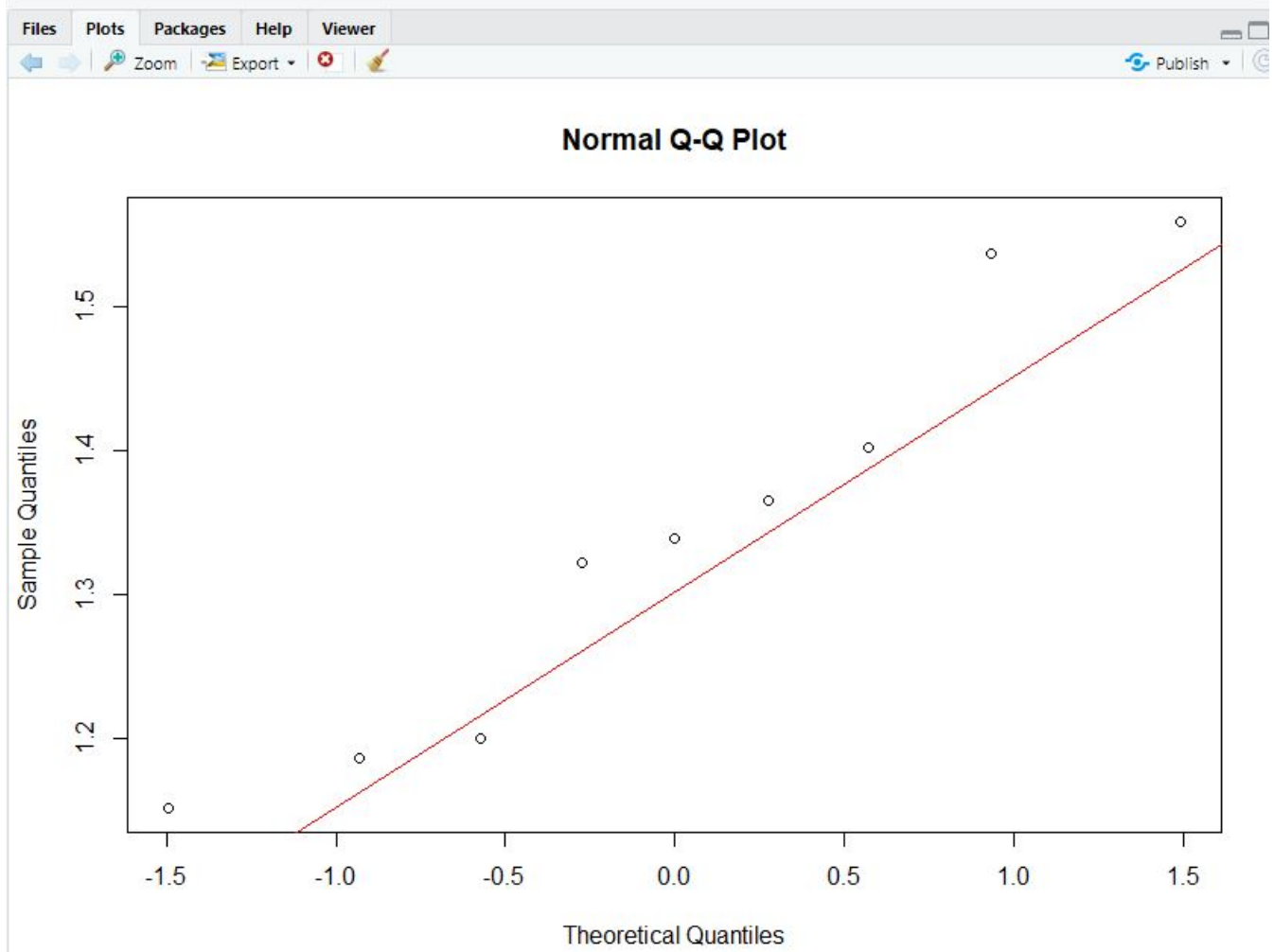
가 된다.

이 값을 각각 구하면,  $0.2739 + -2.306 * 0.135/3$  에서, 0.170, 0.378로 프로그램을 이용하여 나온 결과와 매우 유사함을 확인할 수 있다.

(d) : 가장 간편하게 normality 가정을 확인하는 방법은 qqplot을 이용하여 확인하는 것이다. qqplot을 그리는 과정은, 표준정규분포의 quantile 값들과 실제 데이터의 quantile값들을 각각 x좌표와 y좌표로 가지는 점들을 그래프로 찍어보아, 직선으로 값들이 배열되는 경향을 가지면 데이터가 정규분포를 따른다고 추측할 수 있다.

각각의 데이터가 정규분포를 따르는지 확인하기 위해, r을 이용하여 각 데이터에 따른 qqplot을 그려보았다.

```
> qqnorm(a)
> qqline(a, col = "red")
> |
```

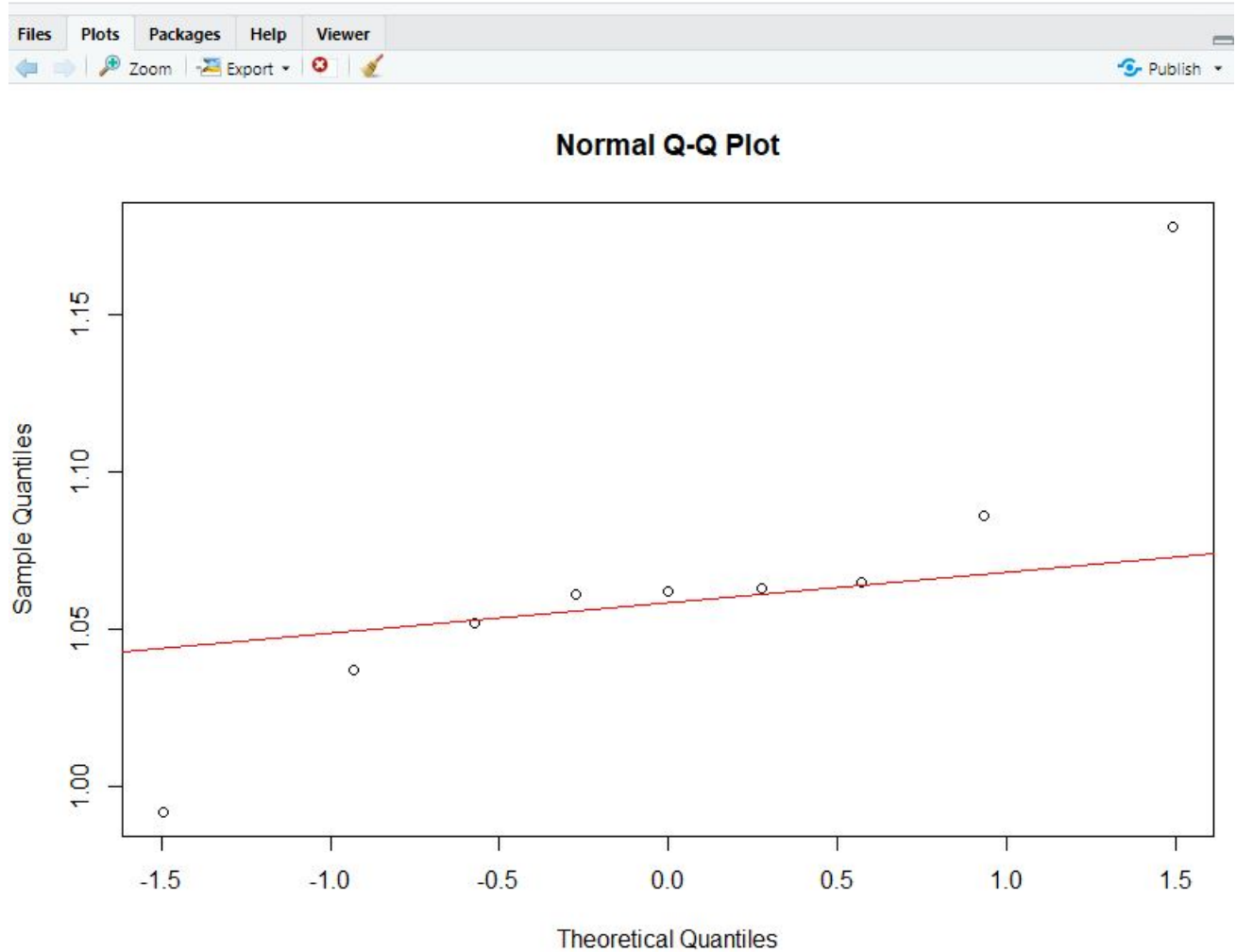


## 실험 계획 및 실습

### Homework #1) 2014-16757 김보창

karlsruhe method에 따른 그래프

```
> qqnorm(b)
> qqline(b, col = "red")
> |
```



lehigh method에 따른 그래프

karlsruhe method는 normality가 만족된다고 생각할 수 있지만,  
lehigh method는 normality가 만족된다고 생각하기에는 애매해서 다음과 같은 test를 추가로 진행하였다.  
표본수가  $n = 2000$ 일때, Shapiro-Wilks test라는 normality를 체크하는 test가 존재하고, R에서 이 test를 지원하기 때문에 각 데이터에 대해 test를 해보았다.

## 실험 계획 및 실습

### Homework #(1) 2014-16757 김보창

---

```
> shapiro.test(a)

      shapiro-wilk normality test

data:  a
W = 0.92905, p-value = 0.4724

> shapiro.test(b)

      shapiro-wilk normality test

data:  b
W = 0.84182, p-value = 0.06051
```

a가 kalsruhe method, b가 lehigh method의 결과이다.

샤피로-윌크스 테스트의  $H_0$ 는 data가 normal distribution을 따름이고,  $H_1$ 은 normal distribution을 따르지 않음이다.

이때,  $\alpha = 0.05$ 로 두었을때, 두경우 모두 p-value가 0.05보다 크기때문에 normality 가정을 충족한다고 생각할 수 있다.

다만, lehigh method의 경우는 p-value의 값이 매우 작은편이기 때문에, normality assumption을 강하게 긍정할 수 있다고는 이야기 할 수 없게된다.

(e): 두 데이터의 차이에 대해서도 위와같이 체크해보았다.

## 실험 계획 및 실습

### Homework #(1) 2014-16757 김보창

```
> qqnorm(d)
> qqline(d, col = "red")
> shapiro.test(d)
```

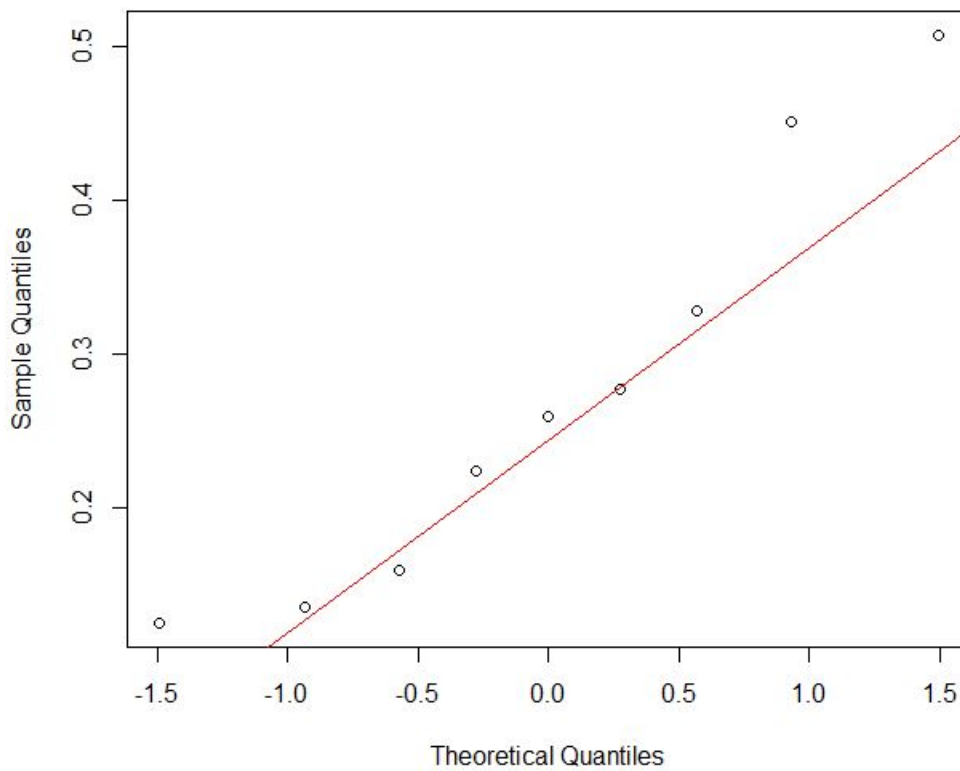
shapiro-wilk normality test

data: d  
W = 0.91678, p-value = 0.3663

```
> |
```



Normal Q-Q Plot



그래프의 형태로 보아서나, p-value의 값으로 보아서나  $\alpha = 0.05$  일때 normality 가정을 충족한다고 생각할 수 있다.

(f) : paired t-test에서의 normality assumption의 역할은, 결국

$$T = \frac{\bar{d} - (u_1 - u_2)}{\frac{S_d}{\sqrt{n}}} \sim t(n-1)$$

위와 같이 test statistic이 t-distribution을 따르게 보장하는 역할을 한다.

## 실험 계획 및 실습

### Homework #1 2014-16757 김보창

이러한 normality assumption은 데이터의 차이  $d_j \stackrel{\text{i.i.d.}}{\sim} N(u_1 - u_2, \sigma^2)$  에서 보장되어야 test statistic이 t-distribution을 따름을 보장할 수 있게된다.

만약, 이러한 normality assumption을 보장할 수 없다면, test statistic의 분포가 무엇인지 알 수 없어 신뢰구간, 가설검정과 같은 통계적 기법을 적용하는데 애로사항이 생기게 되므로, normality assumption의 역할이 매우 중요하다.

### 3 Q3

모집단이 정규분포를 따른다고 가정하자. 즉, 타블렛 1과 타블렛 2의 효과를 각각  $a_i \quad i = 1, 2, \dots, n_1, b_j \quad j = 1, 2, \dots, n_2$  라 하면,

$$a_i \stackrel{\text{i.i.d.}}{\sim} N(u_1, \sigma_1^2)$$

$$b_j \stackrel{\text{i.i.d.}}{\sim} N(u_2, \sigma_2^2)$$

가 성립한다고 가정하자. 이때, 각  $a_i$ 와  $b_j$ 는 서로 독립이다.

$H_0 : 2u_1 - u_2 = 0, H_1 : 2u_1 - u_2 \neq 0$  의 가설에서,

위의 정규성 가정에 의해

$$\bar{a} \sim N(u_1, \frac{\sigma_1^2}{n_1})$$

$$\bar{b} \sim N(u_2, \frac{\sigma_2^2}{n_2})$$

가 성립하게 되고,  $\bar{a}$ 와  $\bar{b}$ 는 서로 독립이므로, 독립인 정규분포를 따르는 확률변수의 합에 대한 성질에 의해

$$2\bar{a} - \bar{b} \sim N(2u_1 - u_2, 4\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

$$Z = \frac{2\bar{a} - \bar{b} - (2u_1 - u_2)}{\sqrt{4\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

가 성립한다.

이때,  $\sigma_1, \sigma_2$ 의 값을 알고있으므로, 따라서 test statistic을 다음과 같이 설정할 수 있다.

$$Z_0 = \frac{2\bar{a} - \bar{b}}{\sqrt{4\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1) \quad \text{when } H_0 \text{ is true}$$

이렇게 설정한 경우, two-sided test를 진행하여  $|Z_0| > z_{\frac{\alpha}{2}}$ 이면  $H_0$ 를 기각, 그렇지 않으면  $H_0$ 를 채택하여 level  $\alpha$  test를 진행할 수 있다.

because weight matrix is created randomly, we think probability of get correct answer is  $\frac{1}{c_0}$ . (1 / category number) so, total loss is mean of each data's loss  $\frac{1}{n} \sum_{j=1}^n \text{Loss}(X_j)$ , and each data's loss is  $\sum_{i=1}^c -L_i \log(S_i)$ , when  $c$  is category number,  $L_i$  is 1 when  $X_j$  is really category  $i$ , else 0.  $S_i$  is softmax score  $\frac{e^{t_i}}{\sum_{k=1}^c e^{t_k}}$  so  $\sum_{i=1}^c -L_i \log(S_i)$  goes to  $-\log(0.1)$  because for all  $k$ ,  $e^{t_k}$  expected equal values and  $\sum_{i=1}^c L_i = 1$  so,  $\frac{1}{n} \sum_{j=1}^n \text{Loss}(X_j)$  expected have  $-\log(0.1)$  value.