homework 02 for design of experiments

106354005 統碩一 余佑駿

As part of a larger experiment, Dale (1992) looked at six samples of a wetland soil undergoing a simulated snowmelt. Three were randomly selected for treatment with a neutral pH snowmelt; the other three got a reduced pH snowmelt. The observed response was the number of Copepoda removed from each microcosm during the first 14 days of snowmelt

Reduced pH			Neutral pH		
256	159	149	54	123	248

Using randomization methods, test ***the null hypothesis that the two treatments have equal average numbers of Copepoda versus a two-sided alternative.***

Answer：

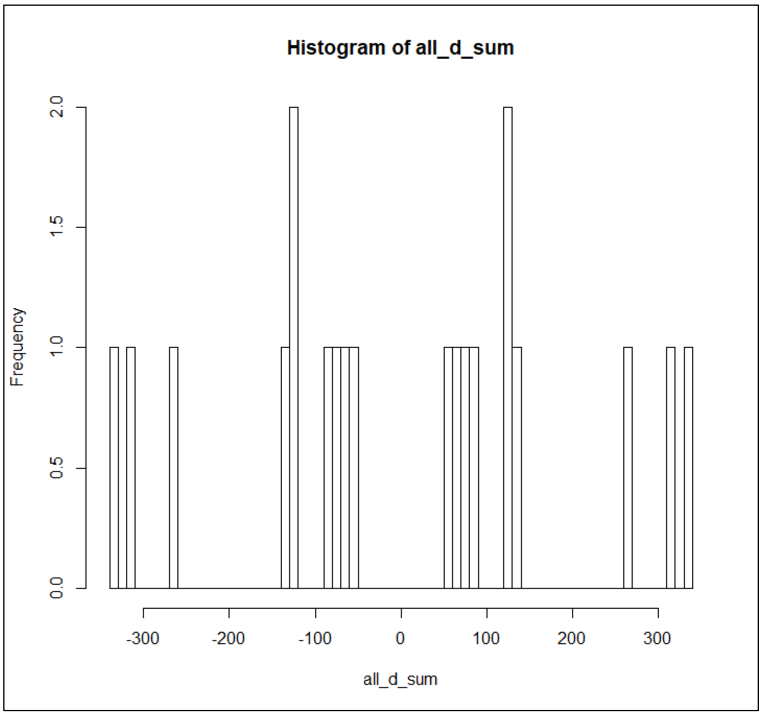
在$null\ \ hypothesis$下，這六個數據可以互換，總共有$C_3^6 = 20$種可能

```
    [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16]
[1,]  256  256  256  256  256  256  256  256  256   256   159   159   159   159   159   159
[2,]  159  159  159  159  149  149  149   54   54   123   149   149   149    54    54   123
[3,]  149   54  123  248   54  123  248  123  248   248    54   123   248   123   248   248
    [,17] [,18] [,19] [,20]
[1,]   149   149   149    54
[2,]    54    54   123   123
[3,]   123   248   248   248
```

接著我們計算每種可能下，每筆資料的差異

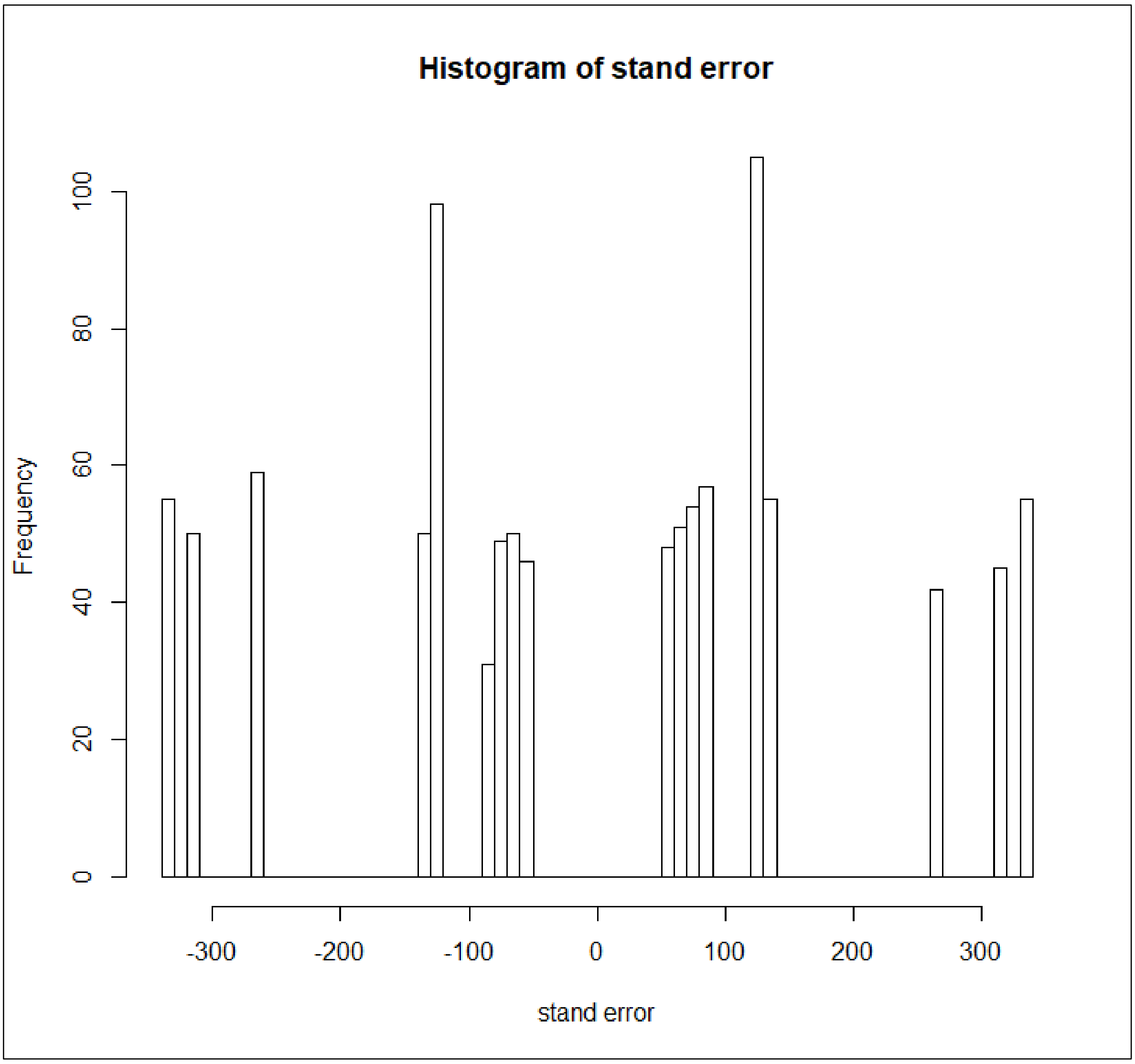
可以畫個直方圖看一下，然後計算雙尾的$p - value = 0.4$

所以不拒絕$null\ \ hypothesis$!!



針對這 1000 筆數據抽 1000 次模擬一下

```
[1] "P hat value : 0.400196"
 [1] "stand error of p hat : 0.000238764"
```



第二題 $Prove\ \ E(\hat{\sigma}^2) = \sigma^2$，其中 $\hat{\sigma}^2 = MSE = \frac{SSE}{N-g} = \frac{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i}.)^2}{N-g}$

Notice that $y_i. = \sum_{j=1}^{n_i} y_{ij}$，$\overline{y_i}. = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i} = treatment\ \ mean$

在不為常態的假設下

$$\mathrm{E(MSE)} = \mathrm{E}\left(\frac{SSE}{N-g}\right) = \frac{1}{N-g}E\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i.\right)^2\right)$$

$$= \frac{1}{N-g}E\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - 2\sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}\bar{y}_i. + \sum_{i=1}^{g}\sum_{j=1}^{n_i}\bar{y}_i.^2\right)$$

$$= \frac{1}{N-g}E\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - 2\sum_{i=1}^{g}\bar{y}..(n_i\overline{y_i.}) + \sum_{i=1}^{g} n_i(\bar{y}_i.)^2\right)$$

$$= \frac{1}{N-g}E\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{g} n_i(\bar{y}_i.)^2\right)$$

$$= \frac{1}{N-g}\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i} E\left[y_{ij}^2\right] - \sum_{i=1}^{g} n_i E[\bar{y}_i.^2]\right)$$

$$= \frac{1}{N-g}\left(\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\sigma^2 + \mu_i^2) - \sum_{i=1}^{g} n_i(\frac{\sigma^2}{n_i} + \mu_i^2)\right)$$

$$= \frac{1}{N-g}\left(N\sigma^2 + \sum_{i=1}^{g} n_i\mu_i^2 - g\sigma^2 - \sum_{i=1}^{g} n_i\mu_i^2\right)$$

$$= \frac{1}{N-g}(N-g)\sigma^2$$

第三題 $Prove\ \ \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{..}\right)^2=\sum_{i=1}^{g}n_i(\bar{y}_{i.}-\bar{y}_{..})^2+\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)^2$

$Solve$:

$$\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{..}\right)^2=\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left[\left(y_{ij}-\bar{y}_{i.}\right)+(\bar{y}_{i.}-\bar{y}_{..})\right]^2$$

$$=\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)^2+\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{y}_{i.}-\bar{y}_{..})^2+2\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)(\bar{y}_{i.}-\bar{y}_{..})$$

$$=\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)^2+\sum_{i=1}^{g}n_i(\bar{y}_{i.}-\bar{y}_{..})^2+2\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)(\bar{y}_{i.}-\bar{y}_{..})$$

Now, just prove the left one term is 0. Then the equation is proved.

Notice that $y_{i.}=\sum_{j=1}^{n_i}y_{ij}$ ， $\overline{y_{i.}}=\frac{\sum_{j=1}^{n_i}y_{ij}}{n_i}=treatment\ \ mean$

$$2\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_{i.}\right)(\bar{y}_{i.}-\bar{y}_{..})=2\sum_{i=1}^{g}\sum_{j=1}^{n_i}y_{ij}\bar{y}_{i.}-2\sum_{i=1}^{g}\sum_{j=1}^{n_i}y_{ij}\bar{y}_{..}-2\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\bar{y}_{i.})^2+2\sum_{i=1}^{g}\sum_{j=1}^{n_i}\bar{y}_{i.}\bar{y}_{..}$$

$$=2\sum_{i=1}^{g}\bar{y}_{i.}(n_i\overline{y_{i.}})-2\sum_{i=1}^{g}\bar{y}_{..}(n_i\overline{y_{i.}})-2\sum_{i=1}^{g}n_i(\bar{y}_{i.})^2+2\sum_{i=1}^{g}n_i(\bar{y}_{i.}\bar{y}_{..})$$

$$=0$$

By the above, the equation is proved. (Q.E.D.)