Design of Experiments

106354010 陳焌彥

2018/3/20

1.

As part of a larger experiment, Dale (1992) looked at six samples of a wetland soil undergoing a simulated snowmelt. Three were randomly se- lected 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.

H_0 :two treatments have equal average numbers of Copepoda

H_1 :not H_0

因為虛無假設兩組平均無差異，故將所有資料以亂數取法分兩組取出，相減除3，接著同樣的方法模擬10000次。

此時檢定統計量為：

$$s = \frac{256 + 159 + 149}{3} - \frac{54 + 123 + 248}{3}$$

接著以檢定統計量求出p-value:

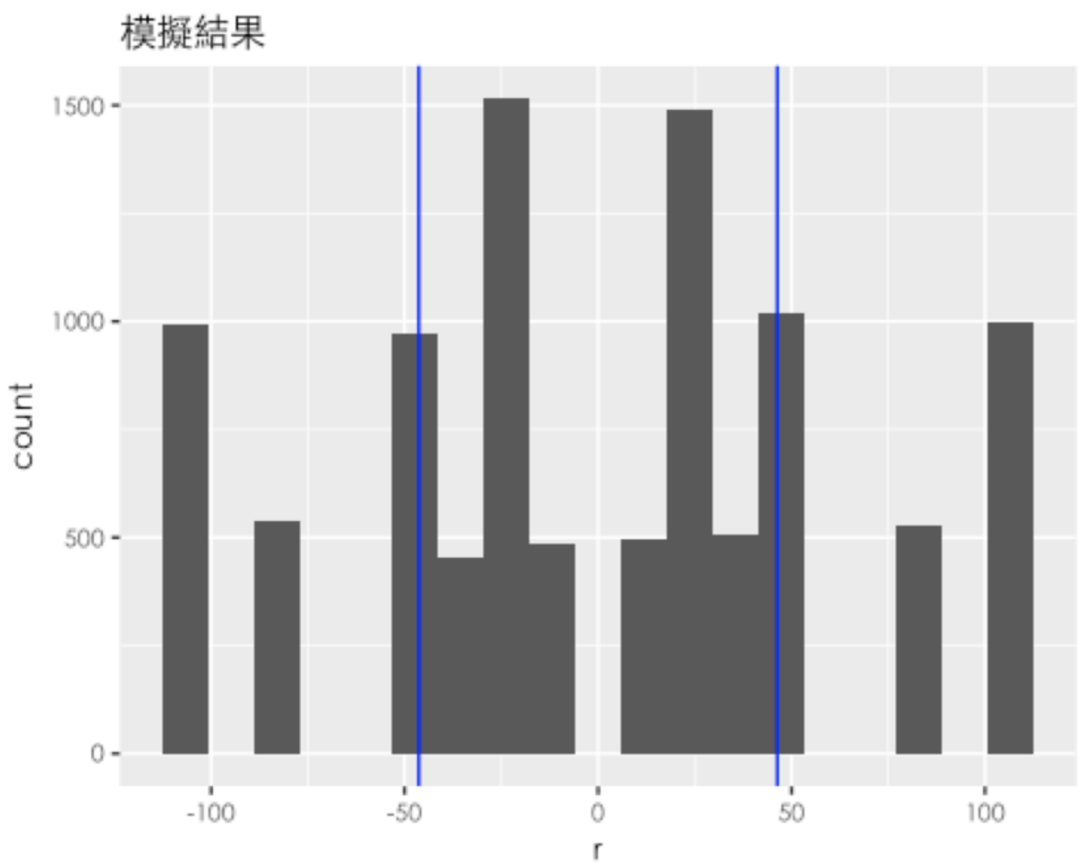
(觀察模擬之資料>=檢定統計量)之個數除以10000

算出p-value為$0.4048 > \alpha = 0.05$

無足夠證據顯示兩組平均有差異，故不拒絕H_0

（Code與圖之結果在下方）

```
ph <- c(256, 159, 149, 54, 123, 248)
s <- (sum(ph[1:3])-sum(ph[4:6]))/3
r <- NULL
for(i in 1:10000){
  ph1 <- sample(ph ,6,F)
  exp <-(sum(ph1[1:3])-sum(ph1[4:6]) )/3
  r <- c(r,exp)
}
ggplot(as.data.frame(r),aes(x = r))+
  geom_histogram(bins = 20)+
  geom_vline(xintercept = c(-46.333333,46.333333), color = "blue")+
  labs(title = "模擬結果",
       x = "r", y = "count")+
  theme(text  = element_text(size = 10 ,family = "STHeitiTC-Light"))
```




```
p_value <- (sum(r<=(-s))+sum(r>=s))/10000

p_value

## [1] 0.4048
```

2.

Derive:

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

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

其中

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

$$\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})=\sum_{i=1}^{ni}y_{ij}-n_i\bar{y}_i=0$$

$$\therefore\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}i.)(\bar{y}_{i.}-\bar{y}\,..)$$

=>

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

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

又

$$\bar{y}_{i.}-\bar{y}_{..}=\alpha_i$$

$$\therefore\underbrace{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}\,..)^2}_{\text{SST}}=\underbrace{\sum_{i=1}^{g}n_i\alpha_i}_{\text{SSTrt}}+\underbrace{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}i.)^2}_{\text{SSE}}$$

3.

$$E(MSE) = E(\frac{SSE}{N-g}) = \frac{1}{N-g}[(\sum_{i=1}^{g}(Y_{ij} - \bar{Y}_{i.})^2]$$

$$= \frac{1}{N-g}\sum_{i=1}^{g}\sum_{j=1}^{n_i} Y_{ij}^2 - 2\sum_{i=1}^{g}\sum_{j=1}^{n_i} Y\bar{Y}_{i.}Y_{ij} + \sum_{i=1}^{g}\sum_{j=1}^{n_i}\bar{Y}_{i.}$$

$$= \frac{1}{N-g}(\sum_{i=1}^{g}\sum_{j=1}^{n_i} E(Y_{ij}) - \sum_{i=1}^{g} n_i\bar{Y}_{i.})$$

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

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

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