Exercise 7.1

Find the smallest sample size giving power of at least .7 when testing equality of six groups at the .05 level when $\zeta = 4n$.

檢定力 $\text{Power} = P(\text{reject } H_0 \mid H_1 \text{ is true}) \geq 0.7$

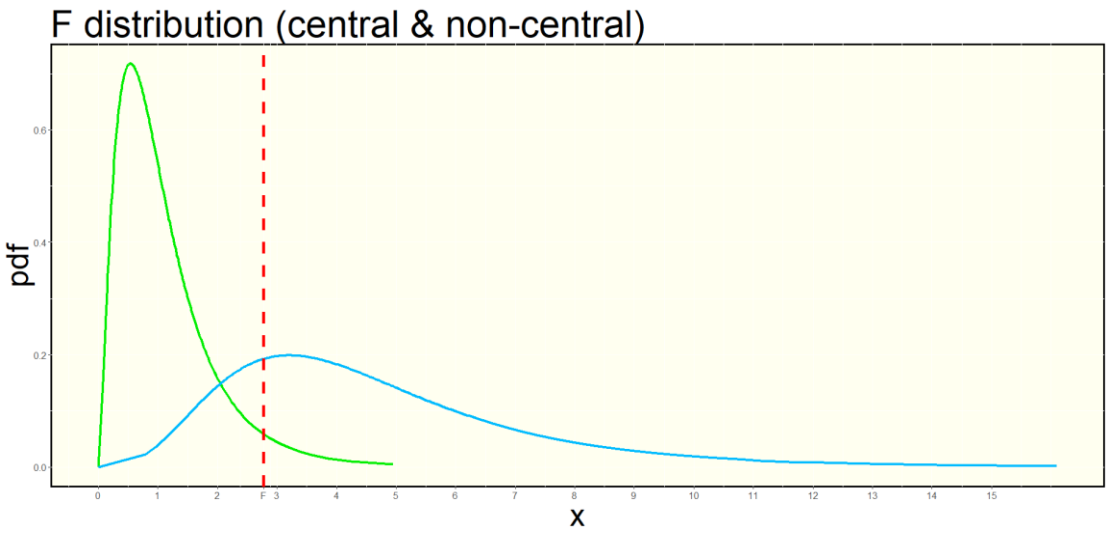
顯著水準 $\varepsilon_i = 0.05$

非中心參數 $\zeta = 4n$

H_0為真時分配為 $F_{(g-1,N-g)}$ (下圖 —綠色線段)
H_1為真時分配為 $F_{(g-1,N-g)}(4n)$ (下圖 —藍色線段)
F 為虛無假設為真下之臨界點 (下圖 —紅色虛線)

我們將 n 帶入計算，並解出下列算式中 n 之最小值

$$P\left(F \geq F_{(g-1,N-g),\epsilon_i} \mid F \sim F_{(g-1,N-g)}(4n)\right) \geq 0.7$$



其中，在固定每組個數相同的情況下，每組個數 n 往上增加一單位，其檢定力也會隨著增加上升，當$n = 4$時，其檢定力為0.7640361超過 0.7，由此得知，此題所需的最小樣本數為$N = 6 \times 4 = 24$。

Exercise 7.3

What is the probability of rejecting the null hypothesis when there are four groups, the sum of the squared treatment effects is 6, the error variance is 3, the group sample sizes are 4, and ε is .01?

每組樣本數為 4，共有 4 組，總共 16 個樣本

非中心參數 $\zeta = \sum_{i=1}^{4}\left(\frac{\alpha_i^*}{\frac{\sigma}{\sqrt{n_i}}}\right)^2 = \sum_{i=1}^{4} n_i\left(\frac{\alpha_i^*}{\sigma}\right)^2 = 8$

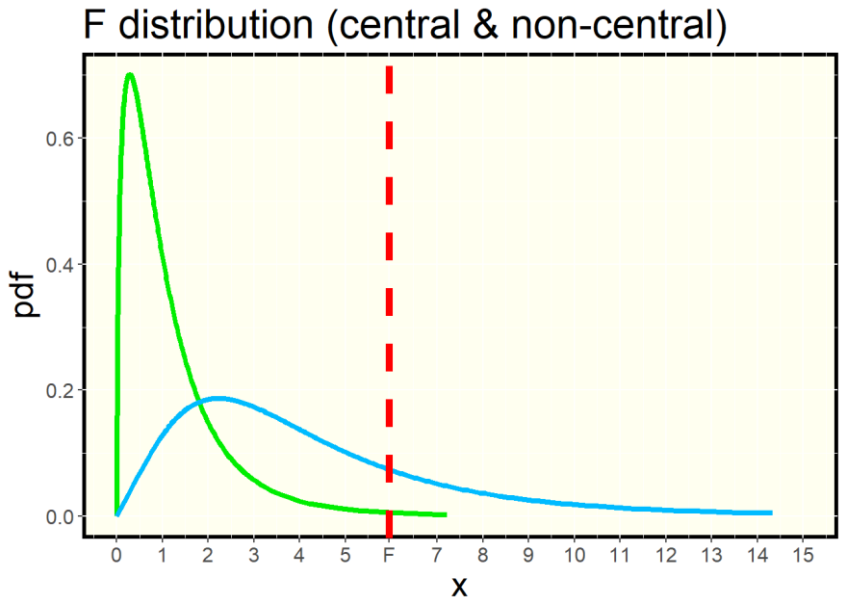
虛無假設下之分配為 $F_{(g-1,N-g)} = F_{(3,12)}$ (下圖 —綠色線段)

對立假設下之分配為 $F_{(g-1,N-g)}(\zeta) = F_{(3,12)}(8)$ (下圖 —藍色線段)

F 為虛無假設為真下之臨界點 (下圖 —紅色虛線)

接著計算其樣本檢定力

$P(\ F \geq F_{(g-1,N-g),\epsilon_i}\ |\ F \sim F_{(g-1,N-g)}(\zeta)\ ) = P(\ F \geq F_{(3,12),0.01}\ |\ F \sim F_{(3,12)}(8)\ )$



最後算出其檢定力為0.2260942

Problem 7.2

Nondigestible carbohydrates can be used in diet foods, but they may have effects on colonic hydrogen production in humans. We want to test to see if inulin, fructooligosaccharide, and lactulose are equivalent in their hydrogen production. Preliminary data suggest that the treatmentmeans could be about 45, 32, and 60 respectively, with the error variance conservatively estimated at 35.
How many subjects do we need to have power .95 for this situation when testing at the ε_I = .01 level?

每組的估計值為 45、32、60，$\sigma^2 = 35$， $\alpha_i = \frac{1}{N}\sum_{i=1}^{N}(y_{1i} - y_{0i})$

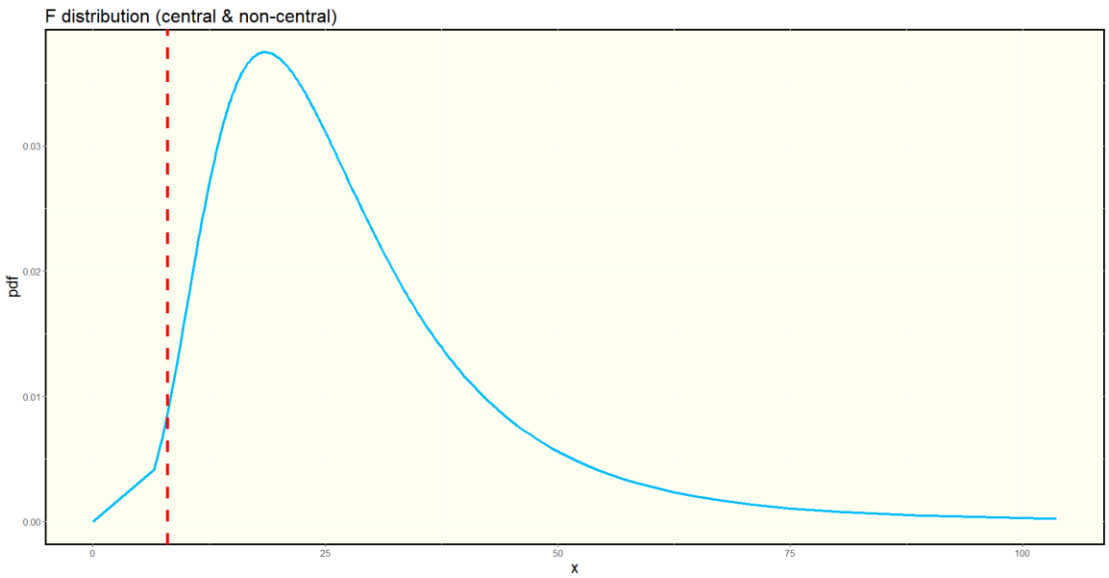
對立假設下分配的非中心參數為 $\zeta^* = \sum_{i=1}^{3}\frac{(\sum_{i=1}^{g}\alpha_i)^2}{\sigma^2/\sqrt{n_i}}$，其中 ζ^* 隨著 n 改變

F 值為 $F_{(g-1,N-g),\epsilon_i} = 8.021517$ (下圖 ─紅色虛線)

將 n 值帶入並計算其檢定力$P(\ F \geq F_{(g-1,N-g),\epsilon_i}\ |\ F \sim F_{(g-1,N-g)}(\zeta^*)\ )$，並找出檢定力$> 0.95$的最小 n 值

經過 R 軟體計算後，我們得到 n=5 時 $\zeta^* = 44.87619$

此時對立假設分配為$F_{(2,12)}$ (下圖 ─藍色線段)



總樣本數為 15，其檢定力大小為0.9808418

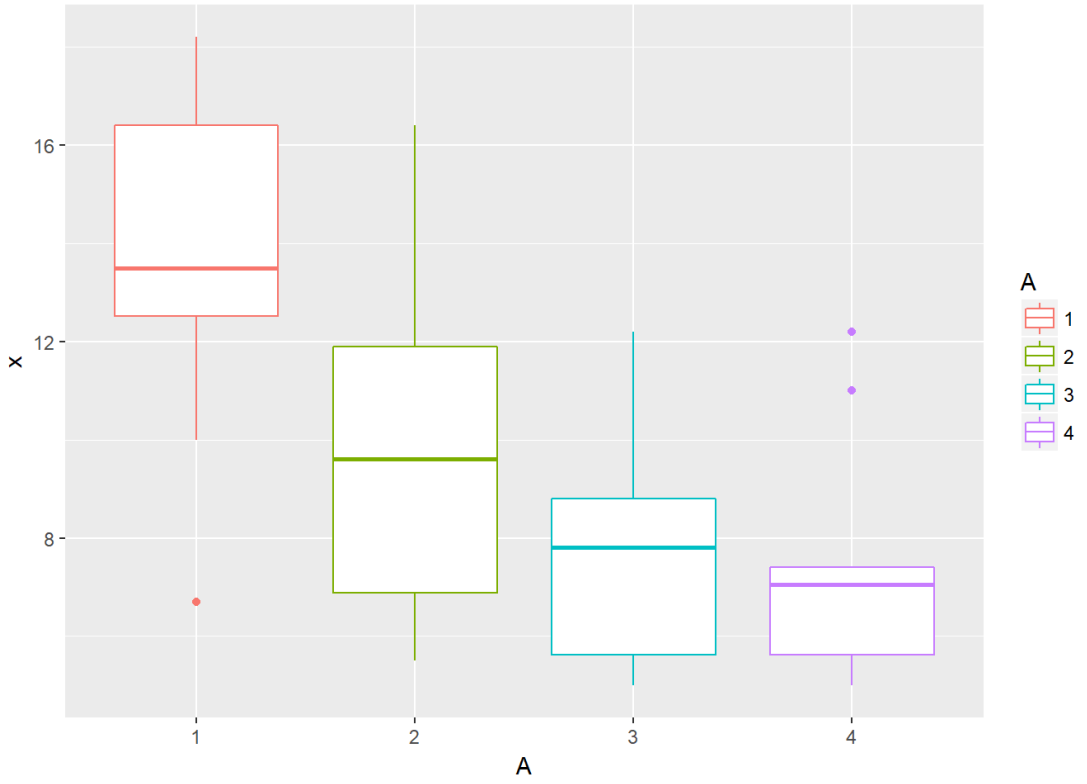
Exercise 6.3

In order to determine the efficacy and lethal dosage of cardiac relaxants,anesthetized guinea pigs are infused with a drug (the treatment) till death occurs.
The total dosage required for death is the response; smaller lethal doses are considered more effective. There are four drugs, and ten guinea pigs are chosen at random for each drug.
Lethal dosages follow.

1	18.2	16.4	10.0	13.5	13.5	6.7	12.2	18.2	13.5	16.4
2	5.5	12.2	11.0	6.7	16.4	8.2	7.4	12.2	6.7	11.0
3	5.5	5.0	8.2	9.0	10.0	6.0	7.4	5.5	12.2	8.2
4	6.0	7.4	12.2	11.0	5.0	7.4	7.4	5.5	6.7	5.5

Test the null hypothesis that the treatments have the same effect.
If they are not, determine which drugs are equivalent,which are more effective and which less effective.

將各組資料繪出盒形圖



並將資料帶入做 ANOVA 檢定

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
組數	3	265.5	88.49	9.865	0.0000691
Residuals	36	322.9	8.97		

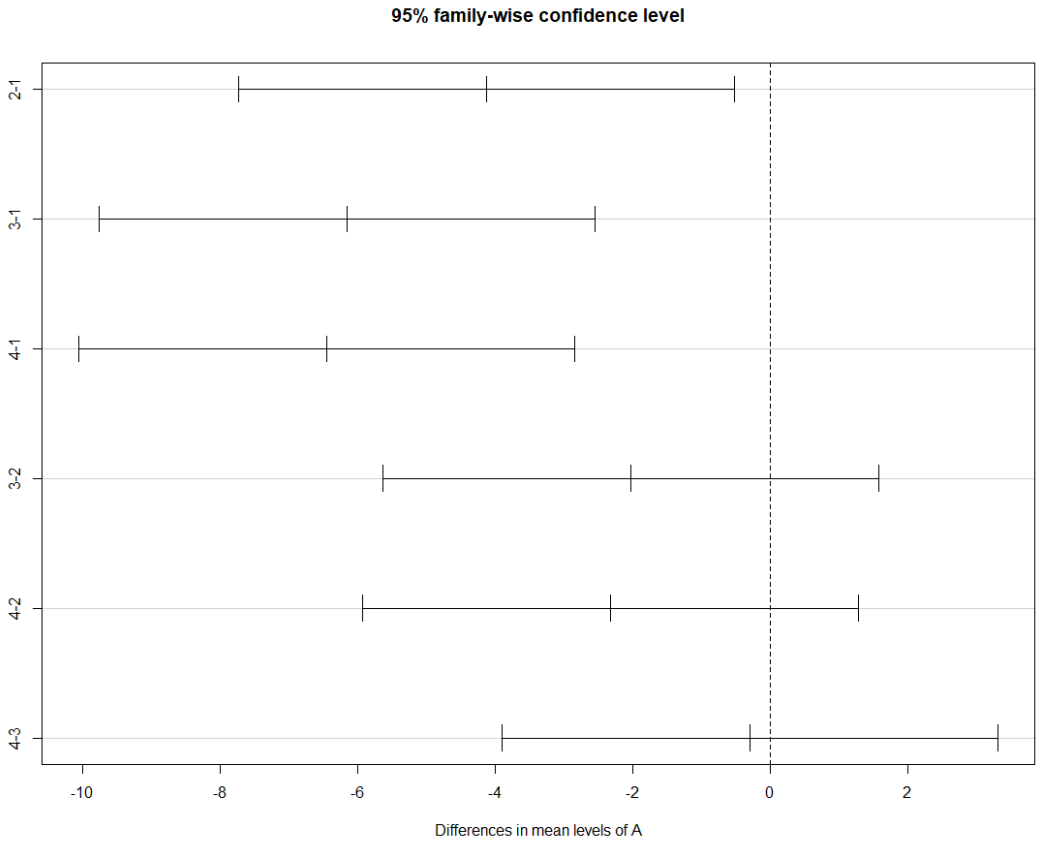
給定顯著水準$\alpha = 0.05$

因為P_value < 0.05，拒絕虛無假設，即各組平均有顯著差異

接下來我們進一步的來討論各組之間的成對平均是否相等
我們使用了 Turcky 多重比較 (檢定兩兩之間是否有顯著差異)

給定顯著水準為 0.05

	diff	lwr	upr	P adj
2-1	-4.13	17.737394	-0.5226057	0.0195604
3-1	-6.16	-9.767394	-2.5526057	0.0002860
4-1	-6.45	-10.057394	-2.8426057	0.0001496
3-2	-2.03	-5.63794	1.5773943	0.4390060
4-2	-2.32	-5.927394	1.2873943	0.3224748
4-3	-0.29	-3.897394	3.3173943	0.9963493



從上圖我們發現第一組與其他組有顯著差異(P_value < 0.05)
我們推論第一組藥品的藥劑量較高，表示其藥效較低

【檢查與驗證 ANOVA 假設】

1. 常態性

這裡我們使用 shapiro test 來檢定資料是否服從常態

```
Shapiro-Wilk normality test

data: residuals

W = 0.97947, p-value = 0.6699
```

由於P_value ＞ 0.05，即資料服從常態分配

2. 獨立性

這裡我們使用 chisq test 來檢定資料是否獨立

```
Pearson's Chi-squared test

data: table(data)

X-squared = 43.6, df = 36, p-value = 0.1796
```

因為 P_value ＞ 0.05，即資料具有獨立性性質

3. 同質性

這裡我們使用 Bartlett 來檢定資料是否具有相同的變異數

```
Bartlett test of homogeneity of variances

data: x by A

Bartlett's K-squared = 2.8347,df = 3, p-value = 0.4178
```

因為 P_value ＞ 0.05，即資料具有變異數同質性質

即此組資料符合 ANOVA 假設前提