这里只列出结果，代码详见hw3.rmd。

Problem 5.7

μ_1, μ_2, μ_3 的 Bonferroni 置信区间 $[3.64, 5.64], [37.10, 53.70], [8.85, 11.08]$,
联合 T^2 置信区间 $[3.40, 5.88], [35.05, 55.75], [8.57, 11.36]$. T^2 置信区间更大。

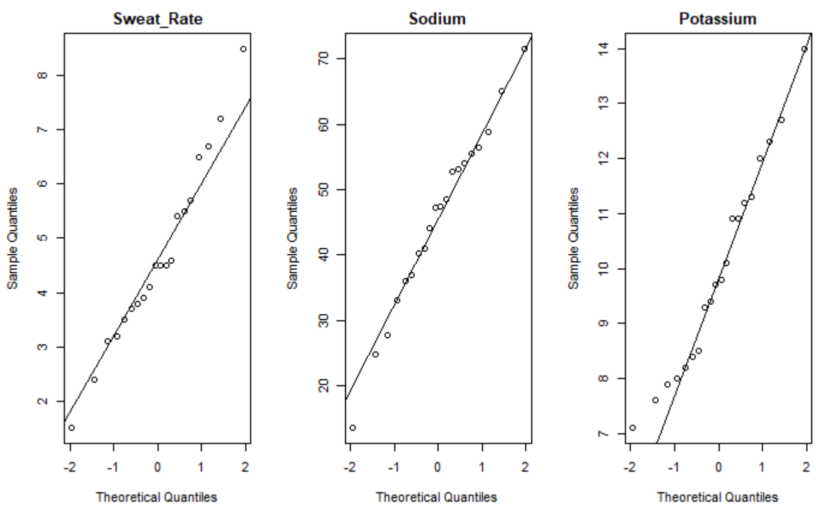

Figure 1: Q-Q plot

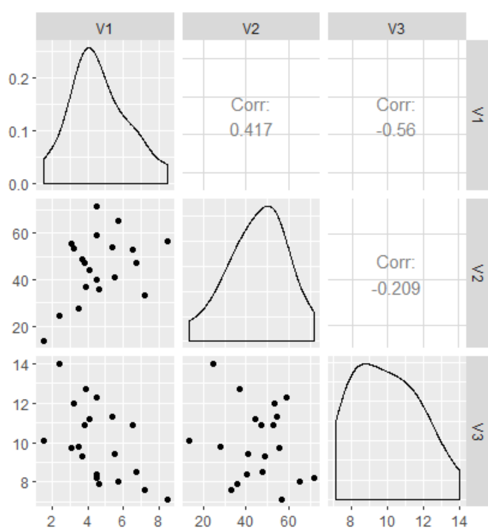

Figure 2: scatter plot matrix

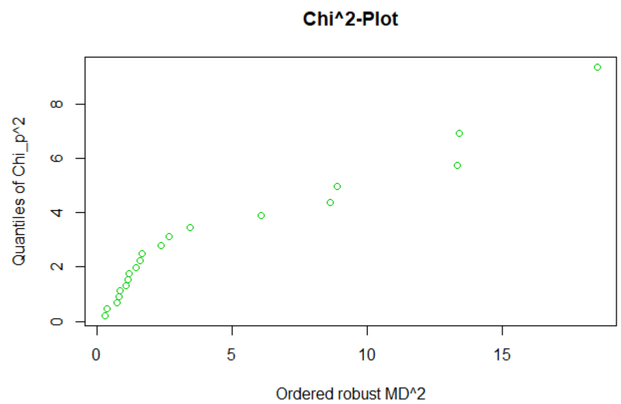

Figure 3: Chi-square plot

由以上各图可知，$x1, x2, x3$ 的分布都是有偏的。正态性假设不合理。

Problem 6 Q1

设想如下两种情形。

(1). 比较两种药物的效果差异。对于一群病人（上百个）中的每一位，比较其服药后一段时间内血液中血细胞含量的变化。这种在同一个病人身上的成对比较是有意义的。

(2). 比较两家工厂生产的元件质量差异。在两家工厂生产的同类元件中，各抽取一百个，统计每个元件的重量，这时成对比较就不合理，因为一对元件之间没有事实上的对应关系。

Problem 6 Q2

(a).$H_0 : \mu_1 - \mu_2 = 0, \alpha = 0.05$, 根据已有数据，

$$
\begin{aligned}
S_{pooled} &= \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2} = \begin{bmatrix} 10963.69 & 21505.42 \\ 21505.42 & 63661.31 \end{bmatrix} \\
T^2 &= (\frac{1}{n_1} + \frac{1}{n_2})^{-1}(\bar{X}_1 - \bar{X}_2)^T S_{pooled}^{-1}(\bar{X}_1 - \bar{X}_2) \\
&= 16.1 \\
c &= \frac{p(n_1+n_2-2)}{n_1+n_2-p-1} F_{p,n_1+n_2-p-1}(\alpha) = 6.2 < T^2
\end{aligned}
$$

拒绝 H_0. 对拒绝 H_0 起关键作用的均值线性组合:

$$
S_{pooled}^{-1}(\bar{X}_1 - \bar{X}_2) = (0.0017, 0.0026)^T
$$

(b).

$$T^2 = (\bar{X}_1 - \bar{X}_2)^T (S_1/n_1 + S_2/n_2)^{-1} (\bar{X}_1 - \bar{X}_2) \quad = 15.7$$
$$c = \chi_p^2(\alpha) = 6.0 < T^2$$

拒绝 H_0. 对拒绝 H_0 起关键作用的均值线性组合:

$$(S_1/n_1 + S_2/n_2)^{-1}(\bar{X}_1 - \bar{X}_2) = (0.041, 0.063)^T$$

(c). 根据 Ch6.6，对 $H_0 : \Sigma_1 = \Sigma_2$ 检验如下:

$$v = \frac{1}{2}p(p+1)(g-1) = 3$$
$$u = \frac{2p^2+3p-1}{6(p+1)(g-1)}[\sum_l \frac{1}{n_l - 1} - \frac{1}{\sum_l (n_l - 1)}] = 0.022$$
$$C = (1-u)\{[\sum_l (n_l - 1)\ln|S_p| - \sum_l [(n_l - 1)]\ln|S_l|]\} = 18.9$$
$$\chi_v^2(\alpha) = 6.0 < C$$

拒绝 H_0.

(d). 既然协方差矩阵相等的假设不成立，(b) 的结果更可信。

Problem 6.11

$$L(\mu_1, \mu_2, \Sigma) = \frac{1}{(2\pi)^{\frac{p(n_1+n_2)}{2}} |\Sigma|^{\frac{n_1+n_2}{2}}}$$
$$\exp\{-\frac{1}{2}[\sum_{i=1}^{n_1}(x_i - \mu_1)^T\Sigma^{-1}(x_i - \mu_1) + \sum_{j=1}^{n_2}(x_j - \mu_2)^T\Sigma^{-1}(x_j - \mu_2)]\}$$
$$l = \log L = -\frac{1}{2}[\sum_{i=1}^{n_1}(x_i - \mu_1)^T\Sigma^{-1}(x_i - \mu_1) + \sum_{j=1}^{n_2}(x_j - \mu_2)^T\Sigma^{-1}(x_j - \mu_2)]$$
$$- \frac{p(n_1+n_2)}{2}\log 2\pi - \frac{n_1+n_2}{2}\log|\Sigma|$$

对 μ_1, μ_2, Σ 分别求导，其中 μ_1, μ_2 的计算过程与单总体情形一致，此处略去过程。结果为 $\hat{\mu_1} = \bar{X}_1, \hat{\mu_2} = \bar{X}_2$.

$$\nabla_\Sigma l = -\frac{n_1+n_2}{2}\Sigma^{-1} + \frac{1}{2}[\sum_{i=1}^{n_1}\Sigma^{-1}(x_i - \mu_1)(x_i - \mu_1)^T\Sigma^{-1} + \sum_{j=1}^{n_2}\Sigma^{-1}(x_j - \mu_2)(x_j - \mu_2)^T\Sigma^{-1}]$$
$$= \frac{\Sigma^{-1}}{2}\{-(n_1+n_2) + [\sum_{i=1}^{n_1}(x_i - \mu_1)(x_i - \mu_1)^T + \sum_{j=1}^{n_2}(x_i - \mu_1)(x_i - \mu_1)^T]\Sigma^{-1}\}$$
$$= \frac{\Sigma^{-1}}{2}\{-(n_1+n_2)] + [(n_1 - 1)S_1 + (n_2 - 2)S_2]\Sigma^{-1}\}$$

令 $\nabla_\Sigma l = 0$, 得 $\hat{\Sigma} = \frac{(n_1-1)S_1 + (n_2-2)S_2}{n_1+n_2}$.