STAT610 - HWK Solution 6

2.2.13 The likelihood is given by

$$
\begin{aligned}
L_X(\theta) &= \prod_{i=1}^{n} I(\theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2}) \\
&= I(\theta - \frac{1}{2} \leq x_{(1)} \leq x_{(n)} \leq \theta + \frac{1}{2}) \\
&= I(X_{(n)} - \frac{1}{2} \leq \theta \leq X_{(1)} + \frac{1}{2})
\end{aligned}
$$

Where $x_{(1)}, x_{(n)}$ are rank statistics. Here, any $\hat{\theta}$ satisfying $X_{(n)} - \frac{1}{2} \leq \hat{\theta} \leq X_{(1)} + \frac{1}{2}$, $L(\hat{\theta}, x)$ takes the maximum.Therefore, any such $\hat{\theta}$ is an MLE.

2.2.15 Since $T(X)$ is sufficient, by factorization theorem

$$
L_X(\theta) = h(X)g[T(X), \theta]
$$

Thus, for any given X,

$$
\hat{\theta} = \arg\min_{\theta} \ L_X(\theta) = \arg\min_{\theta} \ g[T(X), \theta]
$$

which only depends on $T(X)$.

2.2.21 The likelihood is given by

$$
L_X(\mu, \sigma) = \prod_{i=1}^{n} \left[\frac{9}{10\sigma} \varphi \left(\frac{x_i - \mu}{\sigma} \right) + \frac{1}{10} \varphi \left(x_i - \mu \right) \right].
$$

Note that

$$
f(x, \theta) = \begin{cases} \left(\frac{9}{10\sigma} + \frac{1}{10} \right) \varphi(0) \overset{\sigma^2 \to 0}{\longrightarrow} \infty & x = \mu \\ \frac{9}{10\sigma} \varphi \left(\frac{x - \mu}{\sigma} \right) + \frac{1}{10} \varphi(x - \mu) \overset{\sigma^2 \to 0}{\longrightarrow} \frac{1}{10} \varphi(x - \mu) & x \neq \mu \end{cases}
$$

If μ equals to one of $x_1, \cdots, x_n$, say x_1, then

$$
L_X(\mu, \sigma^2) = \left(\frac{9}{10\sigma} + \frac{1}{10} \right) \varphi(0) \prod_{i=2}^{n} \left[\frac{9}{10\sigma} \varphi \left(\frac{x_i - \mu}{\sigma} \right) + \frac{1}{10} \varphi \left(x_i - \mu \right) \right] \to \infty \qquad \text{as } \sigma^2 \to 0
$$

Therefore MLE does not exist and $\sup_{\mu, \sigma^2} L_X(\mu, \sigma^2) = \sup_{\sigma^2} L_X(\hat{\mu} = x_1, \sigma^2) = \infty$.
For any $\breve{\mu} \notin \{x_1, \cdots, x_n\}$,

$$
L_X(\breve{\mu}, \sigma^2) = \prod_{i=1}^{n} \left[\frac{9}{10\sigma} \varphi \left(\frac{x_i - \breve{\mu}}{\sigma} \right) + \frac{1}{10} \varphi \left(x_i - \breve{\mu} \right) \right] \to \begin{cases} \prod_{i=1}^{n} \left[\frac{1}{10} \varphi \left(x_i - \breve{\mu} \right) \right] & \text{as } \sigma^2 \to 0 \\ 0 & \text{as } \sigma^2 \to \infty \end{cases}
$$

Moreover, $L_X(\breve{\mu}, \sigma^2)$ is continuous wrt $\sigma^2 \in (0, +\infty)$, so $\sup_{\sigma^2} L_X(\breve{\mu}, \sigma^2) < \sup_{\mu, \sigma^2} L_X(\mu, \sigma^2) = \infty$.

2.2.24 Solve it by R.

```
feed=c(-1,-1,1,1,-1,-1,1,1,0,0,-sqrt(2),sqrt(2),rep(0,8))
speed=c(rep(-1,4),rep(1,4),-sqrt(2),sqrt(2),rep(0,10))
life=c(54.5,66,11.8,14,5.2,3,.8,.5,86.5,.4,
       20.1,2.9,3.8,2.2,3.2,4,2.8,3.2,4,3.5)
y=log(life)
z1=(feed-13)/6
z2=(speed-900)/300
z11=z1^2
z22=z2^2
z12=z1*z2

fit1=lm(y~z1+z2)  # Model of (a)
fit2=lm(y~z1+z2+z11+z22+z12)   # Model of (b)

### (a) ###
> fit1
Coefficients:
(Intercept)           z1           z2
-1438.720       -4.741     -476.686
>anova(fit1)
Analysis of Variance Table
Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
z1         1  7.4927  7.4927  34.034 1.993e-05 ***
z2         1  30.2973 30.2973 137.619 1.422e-09 ***
Residuals 17  3.7426  0.2202
 ==> The value of contrast function is 3.7426 which is SSE.

### (b) ###
> fit2
Coefficients:
(Intercept)           z1           z2          z11      z22          z12
  231108.05      -332.89    154791.78        15.06     25925.44      -131.14
> anova(fit2)
Analysis of Variance Table
Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
z1         1  7.4927  7.4927  84.8577 2.557e-07 ***
z2         1 30.2973 30.2973 343.1306 3.033e-11 ***
z11        1  1.7398  1.7398  19.7042 0.0005612 ***
z22        1  0.7242  0.7242   8.2016 0.0125035 *
z12        1  0.0425  0.0425   0.4809 0.4993384
Residuals 14  1.2362  0.0883
 ==> The value of contrast function is 1.2362 which is SSE.
```

2.2.30 Without loss of generality that $n_1 = n_2 = \cdots = n_q = 0, n_{q+1} > 0, \cdots, n_k > 0$, then

$$L_X(\theta) = \prod_{j=q+1}^{k} \theta_j^{n_j}$$

which vanishes if $\theta_j = 0$ for any $j = q+1, \cdots, k$. Therefore the MLE $\hat{\theta}_j > 0$ for $j = q+1, \cdots, k$. Since $\sum_{j=1}^{k} \theta_j = 1$,

$$\begin{aligned}
L_X(\theta) &= \left(\prod_{j=q+1}^{k-1} \theta_j^{n_j}\right)\left(1 - \sum_{j=1}^{k-1} \theta_j\right)^{n_k} \\
\Rightarrow l(\theta) &= \log L_X(\theta) = \sum_{j=q+1}^{k-1} n_j \log \theta_j + n_k \log(1 - \sum_{j=1}^{k-1} \theta_j) \\
\Rightarrow \frac{\partial l(\theta)}{\partial \theta_j} &= \begin{cases} \dfrac{-n_k}{1 - \sum_{j=1}^{k-1} \theta_j} = \dfrac{-n_k}{\theta_k} & j = 1, \cdots, q \\ \dfrac{n_i}{\theta_i} - \dfrac{n_k}{1 - \sum_{j=1}^{k-1} \theta_j} = \dfrac{n_i}{\theta_i} - \dfrac{n_k}{\theta_k} & j = q+1, \cdots, k-1 \end{cases}
\end{aligned}$$

For $j = 1, \cdots, q$, $\frac{\partial l(\theta)}{\partial \theta_j} < 0$, likelihood is inceasing as θ_j decreases. So the likelihood is maximized at $\hat{\theta}_1 = \cdots = \hat{\theta}_q = 0$.
For $j = q+1, \cdots, k-1$,

$$\frac{n_j}{\hat{\theta}_j} - \frac{n_k}{\hat{\theta}_k} = 0 \quad \Rightarrow \hat{\theta}_j = \frac{n_j}{n_k}\hat{\theta}_k \quad \Rightarrow \hat{\theta}_j = \frac{n_j}{n}. \qquad (\text{by } \sum_{j=1}^{k} \hat{\theta}_j = \sum_{j=q+1}^{k} \hat{\theta}_j = 1)$$

2.2.39 The likelihood is given by

$$L(\lambda_1, \lambda_2; S, S_1, S_2) = \frac{e^{-n(\lambda_1+\lambda_2)}(n(\lambda_1+\lambda_2))^S}{S!} \frac{e^{-m\lambda_1}(m\lambda_1)^{S_1}}{S_1!} \frac{e^{-m\lambda_2}(m\lambda_2)^{S_2}}{S_2!} \quad (\lambda_1 > 0 \text{ and } \lambda_2 > 0)$$

$$\Rightarrow l = \log L = -(m+n)(\lambda_1+\lambda_2) + S\log(\lambda_1+\lambda_2) + S_1 \log(\lambda_1) + S_2 \log(\lambda_2) + \log(\frac{n^S m^{S_1+S_2}}{S! S_1! S_2!})$$

$$\Rightarrow \begin{cases} \dfrac{\partial l(\lambda_1,\lambda_2)}{\partial \lambda_1} = -(m+n) + \dfrac{S}{\lambda_1+\lambda_2} + \dfrac{S_1}{\lambda_1} = 0 \\ \dfrac{\partial l(\lambda_1,\lambda_2)}{\partial \lambda_2} = -(m+n) + \dfrac{S}{\lambda_1+\lambda_2} + \dfrac{S_2}{\lambda_2} = 0 \end{cases}$$

$$\Rightarrow \hat{\lambda}_1 = \frac{S_1(S+S_1+S_2)}{(m+n)(S_1+S_2)}, \quad \hat{\lambda}_2 = \frac{S_2(S+S_1+S_2)}{(m+n)(S_1+S_2)}, \quad \text{if } S_1 \neq 0 \text{ and } S_2 \neq 0$$

- If $S_1 = 0$ and $S_2 > 0$

$$\frac{\partial l}{\partial \lambda_1} = -(m+n) + \frac{S}{\lambda_1+\lambda_2} \tag{1}$$

$$\frac{\partial l}{\partial \lambda_2} = -(m+n) + \frac{S}{\lambda_1+\lambda_2} + \frac{S_2}{\lambda_2} \tag{2}$$

(1) and (2) can not be 0 simultaneously since (2) is always greater than (1).
From (1), for any given λ_2, the likelihood function is maximized at

$$\hat{\lambda}_1 = \max\{0, S/(m+n) - \lambda_2\}$$

if assuming λ_1 can be 0. Therefore, maximizing the likelihood function is equivalent to maximizing

$$l(\hat{\lambda}_1, \lambda_2) = \begin{cases} -(m+n)(\frac{S}{m+n}) + S\log(\frac{S}{m+n}) + S_2\log(\lambda_2) & \lambda_2 < \frac{S}{m+n} \\ -(m+n)(\lambda_2) + S\log(\lambda_2) + S_2\log(\lambda_2) & \lambda_2 \geq \frac{S}{m+n} \end{cases}.$$

When $\lambda_2 < \frac{S}{m+n}$, $l(\hat{\lambda}_1, \lambda_2)$ increases as λ_2 increases.
When $\lambda_2 \geq \frac{S}{m+n}$, $l(\hat{\lambda}_1, \lambda_2)$ is maximized at $\hat{\lambda}_2 = \frac{S+S_2}{m+n}$.
Thus,

$$\sup_{\lambda_1>0,\lambda_2>0} l(\lambda_1, \lambda_2) = l(0, \tfrac{S+S_2}{m+n})$$

Since λ_1 can not be exactly 0, the MLE doesn't exist.
Similarly the MLE doesn't exist when $S_2 = 0$ and $S_1 > 0$

- If $S_1 = S_2 = 0, S > 0$,

$$\frac{\partial l}{\partial \lambda_1} = \frac{\partial l}{\partial \lambda_2} = -(m+n) + \frac{S}{\lambda_1 + \lambda_2} = 0 \Rightarrow \hat{\lambda}_1 + \hat{\lambda}_2 = \frac{S}{m+n}$$

The MLE could be any values of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ such that $\hat{\lambda}_1 + \hat{\lambda}_2 = \frac{S}{m+n}$.

- If $S_1 = S_2 = S = 0$, $L \to \infty$ as $\lambda_1, \lambda_2 \searrow 0$. So the MLE do not exist.

2.2.40 (a) The likelihood is given by

$$L_X(\theta) = \frac{1}{(\theta_1+\theta_2)^n} \exp\left\{-\frac{1}{\theta_1}\sum_{i=1}^n x_i I(x_i > 0) + \frac{1}{\theta_2}\sum_{i=1}^n x_i I(x_i < 0)\right\}.$$

By factorization theorem, $T_1(X) = \sum_{i=1}^n x_i I(x_i > 0)$ and $T_2(X) = \sum_{i=1}^n (-x_i) I(x_i < 0)$ are sufficient for (θ_1, θ_2).

(b) If $T_1 \neq 0$ and $T_2 \neq 0$ $l(\theta_1, \theta_2) = \log L_X(\theta_1, \theta_2) = -n\log(\theta_1+\theta_2) - T_1/\theta_1 - T_2/\theta_2$

$$\begin{cases} \dfrac{\partial}{\partial \theta_1} l(\theta_1, \theta_2) = -\dfrac{n}{\theta_1+\theta_2} + \dfrac{T_1}{\theta_1^2} = 0 \\ \dfrac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) = -\dfrac{n}{\theta_1+\theta_2} + \dfrac{T_2}{\theta_2^2} = 0 \end{cases} \Rightarrow \begin{cases} \hat{\theta}_1 = \dfrac{T_1 + \sqrt{T_1T_2}}{n} \\ \hat{\theta}_2 = \dfrac{T_2 + \sqrt{T_1T_2}}{n} \end{cases}$$

If $T_1 = 0$, $l(\theta_1, \theta_2) = -n\log(\theta_1+\theta_2) - \frac{T_2}{\theta_2}$. For any fixed θ_2, l is increasing as $\theta_1 \searrow 0$. Therefore, MLE does not exist. Similarly, if $T_2 = 0$, MLE does not exist.

2.3.3 The likelihood is given by

$$\begin{aligned} L_X(\theta) &= \theta_1^{2N_1}\theta_2^{2N_2}\theta_3^{2N_3}(2\theta_1\theta_2)^{N_4}(2\theta_1\theta_3)^{N_5}(2\theta_2\theta_3)^{N_5} \\ &= 2^{N_5+N_4+N_6}\theta_1^{2N_1+N_4+N_5}\theta_2^{2N_2+N_4+N_6}\theta_3^{2N_3+N_5+N_6} \\ &= 2^{N_5+N_4+N_6}\exp\{T_1\log(\theta_1/\theta_3) + T_2\log(\theta_2/\theta_3) + 2n\log\theta_3\}, \end{aligned}$$

where $T_1 = 2N_1 + N_4 + N_5$, $T_2 = 2N_2 + N_4 + N_6$ and $T_3 = 2n - T_1 - T_2$.
Let $\eta_1 = \log(\theta_1/\theta_3)$, $\eta_2 = \log(\theta_2/\theta_3)$, $\theta_3 = 1 - \theta_1 - \theta_2$, $A(\eta_1, \eta_2) = -2n\log\theta_3 = 2n\log(1 + e^{\eta_1} + e^{\eta_2})$
Check the conditions of Theorem 2.3.1

(i) The natural parameter space, $\mathbb{R}^2$, is open.

(ii) $\eta = (\eta_1, \eta_2)$ is identifiable $\Rightarrow$ rank is of k-1.

(iii) If $t_1 + t_2 = 2n$ and $c_1 = c_2 = 1$, then $P(T_1 + T_2 > t_1 + t_2) = P(T_1 + T_2 > 2n) = 0$
If $t_1 = 0$ and $c_1) = -1$, then $P(-T_1 > -t_1) = P(T_1 < 0) = 0$. Similarly if $t_2 = 0$.
Thus, when $t_1 + t_2 = 2n$ or $t_1 = 0$ or $t_2 = 0$, MLE does not exist. Otherwise, MLE exists.

Solve $\dot{A}(\eta) = T$ to obtain the MLE,

$$\begin{cases} \dfrac{\partial A(\eta_1, \eta_2)}{\partial \eta_1} = \dfrac{2ne^{\eta_1}}{1 + e^{\eta_1} + e^{\eta_2}} = 2n\theta_1 = T_1 \\ \dfrac{\partial A(\eta_1, \eta_2)}{\partial \eta_2} = \dfrac{2ne^{\eta_2}}{1 + e^{\eta_1} + e^{\eta_2}} = 2n\theta_2 = T_2 \end{cases}$$

So $\hat{\theta}_1 = \frac{T_1}{2n}$, $\hat{\theta}_2 = \frac{T_2}{2n}$ and $\hat{\theta}_3 = 1 - \hat{\theta}_1 - \hat{\theta}_2$.

2.3.10 The multinomial distribution is a canonical exponential family distribution with an open parameter space and rank of $k-1$. By Corollary 2.3.1, MLE exists $\Leftrightarrow (t_1, \cdots, t_{k-1}) \in C_T^0$, where C_T^0 is the interior of the convex support set $C_T = \left\{(t_1, \cdots, t_{k-1}); 0 \le t_j \le n, j = 1, \cdots, k-1, \sum_{j=1}^{k-1} t_j \le n\right\}$.

$$C_T^0 = \{(t_1, \cdots, t_{k-1}); t_j > 0, 1 \le j \le k-1, \sum_{j=1}^{k-1} t_j < n\}$$

Therefore, MLE exists $\Leftrightarrow (T_1, T_2, \cdots, T_{k-1}) \in C_T^0 \Leftrightarrow \forall T_j > 0, 1 \le j \le k$.