Question 1

$Y_1 \cdots Y_n \sim N(\mu, \sigma^2)$

$R_1 = Y_i - Y_1$

$R_2 = Y_i - \bar{Y}$ $\quad \bar{Y} = \sum_{i=1}^{n}$

$\Sigma_i = \sigma^2 I$ is $n \times n$

X is 1_n is $n \times 1$

$R_{n\times 1}$

$$\ell_{REML}(\theta, R) = -\frac{1}{2}\log|\Sigma| - \frac{1}{2}\log|X^T\Sigma^{-1}X| - \frac{1}{2}Q(R, \Sigma, X)$$

$$X^T\Sigma^{-1}X = [1 \cdots 1]\frac{1}{\sigma^2}I\begin{bmatrix}1\\ \vdots \\ 1\end{bmatrix} = \frac{n}{\sigma^2}$$

$$\Rightarrow \ell_{REML}(\theta, R) = -\frac{1}{2}\log(\sigma^2 I) - \frac{1}{2}\log\left(\frac{n}{\sigma^2}\right) - \frac{1}{2}Q(R, \Sigma, X)$$

$$Q(R, \Sigma, X) = R^T\{\Sigma^{-1} - \Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}\}R$$

$$= \{R^T\Sigma^{-1} - R^T\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}\}R$$

$$= R^T\Sigma^{-1}R - R^T\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}R$$

$$= R^T\frac{1}{\sigma^2}IR - R^T\frac{1}{\sigma^2}IX\cdot\left(\frac{\sigma^2}{n}\right)X^T\left(\frac{1}{\sigma^2}\right)R$$

$$= \frac{1}{\sigma^2}R^TR - \frac{1}{\sigma^2}\frac{\sigma^2}{n}\frac{1}{\sigma^2}R^TXX^TR$$

$$= \frac{1}{\sigma^2}R^TR - \frac{1}{\sigma^2 n}R^TJR = \frac{1}{\sigma^2}\sum_{i=1}^{n}R_i^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n}R_i\right)^2$$

$$\ell_{REML}(\theta, R) = -\frac{1}{2}\log(\sigma^2 I) - \frac{1}{2}\log\left(\frac{n}{\sigma^2}\right) - \frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}R_i^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n}R_i\right)^2\right]$$

$\Rightarrow$ if R's vary the $Q(R, \Sigma, X)$ changes

For $R_1 = Y_i - Y_1$

$$\ell_{REML}(\theta, R_1) = -\frac{1}{2}\log(\sigma^2 I) - \frac{1}{2}\log\left(\frac{n}{\sigma^2}\right) - \frac{1}{2}\underbrace{\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - Y_1)^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n}Y_i - Y_1\right)^2\right]}$$

$$= \sum_{i=1}^{n}(Y_i - Y_1)(Y_i - Y_1) = \sum_{i=1}^{n}Y_i^2 - 2Y_iY_1 + Y_1^2$$

$$= \sum_{i=1}^{n}Y_i^2 - 2\sum_{i=1}^{n}Y_iY_1 + \sum_{i=1}^{n}Y_1^2$$

$$\sum_{i=1}^{n}Y_i^2 - 2Y_1\sum_{i=1}^{n}Y_i + nY_1^2$$

$$\left(\sum_{i=1}^{n}Y_i - \sum_{i=1}^{n}Y_1\right)^2$$

$$\left(\sum_{i=1}^{n}Y_i\right)^2 - 2\sum_{i=1}^{n}Y_i\sum_{i=1}^{n}Y_1 + \left(\sum_{i=1}^{n}Y_1\right)^2$$

$$\left(\sum_{i=1}^{n}Y_i\right)^2 - 2Y_1 n\sum_{i=1}^{n}Y_i + (nY_1)^2$$

$$Q(R_1, \Sigma, x) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} Y_i^2 - 2Y_1 \sum_{i=1}^{n} Y_i + nY_1^2\right) - \frac{1}{\sigma^2 n}\left[\left(\sum_{i=1}^{n} Y_i\right)^2 - 2Y_1 n \sum_{i=1}^{n} Y_i + (nY_1)^2\right]$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 - \cancel{\frac{2}{\sigma^2} Y_1 \sum_{i=1}^{n} Y_i} + \cancel{\frac{n}{\sigma^2} Y_1^2} - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i\right)^2 + \cancel{\frac{2}{\sigma^2} Y_1 \sum_{i=1}^{n} Y_i} - \cancel{\frac{n}{\sigma^2} Y_i^2}$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i\right)^2$$

✓

$R_2 = Y_i - \bar{Y}$

$$Q(R_2, \Sigma, X) = \frac{1}{\sigma^2}\sum_{i=1}^{n} R_i^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} R_i\right)^2$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} (Y_i - \bar{Y})^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i - \bar{Y}\right)^2$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} (Y_i^2 - 2\bar{Y} Y_i + \bar{Y}^2) - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \bar{Y}\right)^2$$

$$= \frac{1}{\sigma^2}\left[\sum_{i=1}^{n} Y_i^2 - 2\bar{Y}\sum_{i=1}^{n} Y_i + n\bar{Y}^2\right] - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i - n\bar{Y}\right)^2$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 - \frac{2}{\sigma^2}\bar{Y}\sum_{i=1}^{n} Y_i + \frac{n}{\sigma^2}\bar{Y}^2 - \frac{1}{\sigma^2 n}\left[\left(\sum_{i=1}^{n} Y_i\right)^2 - 2n\bar{Y}\sum_{i=1}^{n} Y_i + n^2\bar{Y}^2\right)$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 - \cancel{\frac{2}{\sigma^2}\bar{Y}\sum_{i=1}^{n} Y_i} + \cancel{\frac{n}{\sigma^2}\bar{Y}^2} - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i\right)^2 + \cancel{\frac{2}{\sigma^2}\bar{Y}\sum_{i=1}^{n} Y_i} - \cancel{\frac{n}{\sigma^2}\bar{Y}^2}$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 - \frac{1}{\sigma^2 n}\left(\sum_{i=1}^{n} Y_i\right)^2$$

✓

Conclusion: The ℓ_{REML} for R_1 is the same as the ℓ_{REML} for R_2.

✓

Question 2

$$\begin{bmatrix}\mu_1\\ \mu_2\\ \mu_3\\ \mu_4\end{bmatrix} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\beta_0 + \begin{bmatrix}t_1\\t_2\\t_3\\t_4\end{bmatrix}\beta_1 + \begin{bmatrix}t_1^2\\t_2^2\\t_3^2\\t_4^2\end{bmatrix}\beta_2 + \begin{bmatrix}t_1^3\\t_2^3\\t_3^3\\t_4^3\end{bmatrix}\beta_3$$

pg 44

$\bar{y} \cdot P$ = orthogonal polynomial estimate

$$\underbrace{\begin{bmatrix}0.5 & 0.5 & 0.5 & 0.5\\ -0.67 & -0.22 & 0.22 & 0.67\\ 0.5 & -0.5 & -0.5 & 0.5\\ -0.22 & 0.67 & -0.67 & 0.22\end{bmatrix}}_{P} \begin{bmatrix}1.14\\2.54\\2.99\\3.47\end{bmatrix} = \begin{bmatrix}5.67\\1.66\\-0.46\\0.22\end{bmatrix}$$

$$\begin{bmatrix}\bar{Y}_1\\ \bar{Y}_2\\ \bar{Y}_3\\ \bar{Y}_4\end{bmatrix} \quad \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$$

→ the second row of P is the contrast of linear component (β_1) $[-0.67 \ -0.22 \ 0.22 \ 0.67] = 0$

→ the third row of P is the contrast of qudratic component (β_2)

→ the fourth row of P is the contrast of the cubic component (β_3)

(a) if the true population linear contrast = 0 it means the linear component has no effect on the overall mear

$P_2 = [-0.67 \quad -0.22 \quad 0.22 \quad 0.67] \div 0.22 = [-3 \quad -1 \quad 1 \quad 3]$

$$[-3 \ -1 \ 1 \ 3]\begin{bmatrix}\bar{Y}_1\\ \bar{Y}_2\\ \bar{Y}_3\\ \bar{Y}_4\end{bmatrix} = -3\bar{Y} - \bar{Y}_2 + \bar{Y}_3 + 3\bar{Y}_4 = \bar{Y}_3 + 3\bar{Y}_4 - [3\bar{Y}_1 + \bar{Y}_2]$$

The difference between the average mean at time 3 and three times the mean at time 4, and three times the mean at time 1 and time 2.

✓

(B) what does it mean to say the quadratic contrast is zero?

if contrast = 0 ⟹ quadratic component has no effect on overal mean.

$[0.5 \quad -0.5 \quad -0.5 \quad 0.5] \div 0.5 = [1 \quad -1 \quad -1 \quad 1]$

$$[1 \; -1 \; -1 \; 1] \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \bar{Y}_3 \\ \bar{Y}_4 \end{bmatrix} = \bar{Y}_1 - \bar{Y}_2 - \bar{Y}_3 + \bar{Y}_4$$

$$= (\bar{Y}_1 + \bar{Y}_4) - (Y_3 + Y_2)$$

⟹ The difference between the average of the means of time 1 and time 4, and the average of the means at time 2 and time 3.

✓

Question 3

(a)

M1 - covariance is unstructured and different for each group

- different means for each level of group and time.
- 8 different combinations of group and time that all have different intercepts,

M2 - covariance is unstructured and not different for each group.

- the mean structure is the same as model 1.

M3 - the covariance is unstructured and not different for each group.

- the model assumes that the marginal mean follow a linear trend with timex.
- each treatment group has its own line, one for placebo and one for active w/ two different slopes and two different intercepts,

M4 - covariance is unstructured and not different for each group.

- the model assumes that the marginal mean follows a linear trend with timex.
- two different slopes (placebo and active) and the same intercept.

M5 - covariance is unstructured and not different for each group.

- 4 different combinations of group (placebo and active) and time14e (baseline and not baseline) that have different intercepts

M6 - covariance is unstructured and the same for both groups.
- model takes t146 and active
- the intercept is the same for both active and placebo.

M7 - covariance is unstructured and the same for both groups.
- model uses different time points and active.
- the intercept (baseline) is the same for both treatments.
- means are different for each timepoint after baseline (t1, t4, t6) for each trt group.

B) • K=100
• seperated by group and time.
• $\mu_i = (\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4})^T$ for the four different time point,

$X_{ij} = (1, t1_{ij}, t4_{ij}, t6_{ij}, t1*active_{ij}, t4*active_{ij}, t6*active_{ij})$

$i = 1, \dots, K \quad j = 1, 2, 3, 4$

where
• $t1 = 1$ when $j = 2$ $\quad t_1 = 0$ otw
• $t4 = 1$ when $j = 3$ $\quad t_2 = 0$ otw
ect.

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_4 & t_6 & t_1{*}\text{active} & t_4{*}\text{active} & t_6{*}\text{active} \\ 1 & t_1 & t_4 & t_6 & t_1{*}\text{active} & t_4{*}\text{active} & t_6{*}\text{active} \\ 1 & t_1 & t_4 & t_6 & t_1{*}\text{active} & t_4{*}\text{active} & t_6{*}\text{active} \\ 1 & t_1 & t_4 & t_6 & t_1{*}\text{active} & t_4{*}\text{active} & t_6{*}\text{active} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{bmatrix}$$

Covariance structure? -2

(c)

t_6

"β_3" is the mean incremental increase of lead blood levels at week 6 compare to baseline for placebo.

t_6* active

"β_6" is the mean incremental increase of interaction effect of week 6 and active. -2

i.e. difference btwn. placebo & active.

(d) To see the treatment effect you would look at the interaction parameters (t_1 * active, t_4 * active, t_6 * active). This would show the treatment effect because its adding an additional incremental increase solely becaus of the active treatment.