Biostat653 Homework 2

Due Wednesday October 11th, 3:10pm, in class.
Name: Xubo YUE (21819315)

Problem 1

Consider the general linear model $y = X\beta + \epsilon, E(\epsilon) = 0, V(\epsilon) = \Sigma$, where y is an n-vector of responses, X is an n by p matrix of covariates, ϵ is an n-vector of residual errors, and Σ is a known n by n symmetric positive definite matrix. Let W be an arbitrary, symmetric positive definite matrix. Let $\hat{\beta}_W$ denote a solution to the following optimization problem

$$\min_{\beta}(y - X\beta)^T W(y - X\beta)$$

The weighted/generalized least square estimator $\hat{\beta}_{\Sigma^{-1}}$ corresponds to $W = \Sigma^{-1}$. Let c be a p-vector (i.e. contrast matrix) and $c^T\beta$ be the corresponding transformed parameter.

a Show that $E(c^T\hat{\beta}_W) = E(c^T\hat{\beta}_{\Sigma^{-1}})$.
Proof: The $\hat{\beta}_W$ can be estimated as

$$\hat{\beta}_W = (X^T W X)^{-1} X^T W Y$$

for the arbitrary (but known) W. So

$$\mathbb{E}[\hat{\beta}_W] = (X^T W X)^{-1} X^T W \mathbb{E}[Y] = (X^T W X)^{-1} X^T W X\beta = \beta$$
$$\Rightarrow \mathbb{E}[c^T\hat{\beta}_W] = c^T\beta$$

Also, we have

$$\mathbb{E}[\hat{\beta}_{\Sigma^{-1}}] = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}X\beta = \beta$$
$$\Rightarrow \mathbb{E}[c^T\hat{\beta}_{\Sigma^{-1}}] = c^T\beta$$

Thus, we proved $E(c^T\hat{\beta}_W) = E(c^T\hat{\beta}_{\Sigma^{-1}})$. ■

b Show that $V(c^T\hat{\beta}_W) \geq V(c^T\hat{\beta}_{\Sigma^{-1}})$.
Proof: The variance of $\hat{\beta}_W$ is

$$V(\hat{\beta}_W) = (X^T W X)^{-1} X^T W \Sigma W X (X^T W X)^{-1}$$

The variance of $\hat{\beta}_{\Sigma^{-1}}$ is

$$V(\hat{\beta}_{\Sigma^{-1}}) = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}\Sigma\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1} = (X^T\Sigma^{-1}X)^{-1}$$

We compare these two variance by substraction

$$\begin{aligned}
&V(c^T\hat{\beta}_W) - V(c^T\hat{\beta}_{\Sigma^{-1}}) \\
&= c^T((X^TWX)^{-1}X^TW - (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1})\Sigma \\
&((X^TWX)^{-1}X^TW - (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1})c \\
&= C_{new}^T\Sigma C_{new} \geq 0
\end{aligned}$$

The last inequality holds because Σ is positive definite.
So we proved $V(c^T\hat{\beta}_W) \geq V(c^T\hat{\beta}_{\Sigma^{-1}})$. ■

Note that, above, for simplicity, we effectively set the sample size N=1. Also, when M=I, the corresponding estimator $\hat{\beta}_I$ (by setting $W = I^{-1}$) is the ordinal least squares estimator.

Problem 2

We simulate a data set with N=1,000 individuals. We assume that for each individual we obtain n=2 repeated measurements. These measurements are simulated from a multivariate normal distribution with mean $(0, 0)^T$ and a covariance matrix $\begin{bmatrix} 2.0 & 0.5 \\ 0.5 & 1.0 \end{bmatrix}$. Our goal is to use the general linear model to fit the data and obtain parameter estimates. In particular, we consider the model $Y_i = 1_2\mu + \epsilon_i$, for i=1,...,N, where Y_i is a 2-vector of outcomes, 1_2 is a 2-vector of 1's, μ is the intercept, ϵ_i is a 2-vector of residual errors. Note that we do not use the notation X here.

(1) Write down code and simulate a data.
Solution: We simulate data by the following code:

```
set.seed(1)
N <- 1000
n <- 2
require(MASS)
dataset <- mvrnorm(N, rep(0, n), matrix(c(2, 0.5, 0.5, 1), n, n))
```

(2) Under a normality assumption that $\epsilon_i \sim MVN(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma)$, write down the log-likelihood and the MLE algorithm to estimate μ and Σ. How do you compute $V(\hat{\mu}_{MLE})$?
Solution: We have the density function

$$f(Y_i) = (2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}}exp\{-\frac{1}{2}(Y_i - 1_2\mu)^T\Sigma^{-1}(Y_i - 1_2\mu)\} \text{ for } i = 1, \ldots, N$$

So we can obtain the likelihood function as

$$L = \prod_{i=1}^{N}(2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}}exp\{-\frac{1}{2}(Y_i - 1_2\mu)^T\Sigma^{-1}(Y_i - 1_2\mu))\}$$

$$l = \sum_{i=1}^{N}\{-\frac{n}{2}log(2\pi) - \frac{1}{2}log|\Sigma| - \frac{1}{2}(Y_i - 1_2\mu)^T\Sigma^{-1}(Y_i - 1_2\mu)\}$$

$$= -\frac{Nn}{2}log(2\pi) - \frac{N}{2}log|\Sigma| - \sum_{i=1}^{N}\frac{1}{2}(Y_i^T\Sigma^{-1}Y_i - Y_i^T\Sigma^{-1}1_2\mu - \mu^T1_2^T\Sigma^{-1}Y_i$$

$$+ \mu^T1_2^T\Sigma^{-1}1_2\mu)$$

By taking partial derivative with respect to μ, we have

$$\frac{\partial l}{\partial \mu} = -\sum_{i=1}^{N}(-1_2^T\Sigma^{-1}Y_i + \mu^T1_2^T\Sigma^{-1}1_2) = 0$$

$$\Rightarrow \hat{\mu} = (\sum_{i=1}^{N}1_2^T\Sigma^{-1}1_2)^{-1}(\sum_{i=1}^{N}1_2^T\Sigma^{-1}Y_i)$$

Again, by taking partial derivative with respect to Σ^{-1}, we have

$$\frac{\partial l}{\partial \Sigma^{-1}} = \frac{N}{2}\Sigma - \sum_{i=1}^{N}\frac{1}{2}(Y_i - 1_2\mu)(Y_i - 1_2\mu)^T = 0$$

$$\Rightarrow \hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{1000}(Y_i - 1_2\hat{\mu})(Y_i - 1_2\hat{\mu})^T$$

The algorithm is summarized as follow:

Algorithm 1 Estimation Algorithm

1: Given $Y_1, Y_2, \ldots, Y_{1000}$. Set $\hat{\Sigma}_0 = I$.
2: for t in 1:T:
3: $\quad \hat{\mu}_{(t)} = (\sum_{i=1}^{N}1_2^T\hat{\Sigma}_{(t-1)}^{-1}1_2)^{-1}(\sum_{i=1}^{N}1_2^T\hat{\Sigma}_{(t-1)}^{-1}Y_i)$.
4: $\quad \hat{\Sigma}_{(t)} = \frac{1}{1000}\sum_{i=1}^{1000}(Y_i - \hat{\mu}'_{(t)})(Y_i - \hat{\mu}'_{(t)})^T$.
5: Until convergence.

To compute $V(\hat{\mu}_{MLE})$, we start from Fisher information matrix

$$\frac{\partial^2 l}{\partial \mu'^2} = -\sum_{i=1}^{N}(1_2^T \Sigma^{-1} 1_2)$$

$$\Rightarrow I(\mu') = (\sum_{i=1}^{N} 1_2^T \Sigma^{-1} 1_2)$$

$$\Rightarrow V(\hat{\mu}_{MLE}) = (\sum_{i=1}^{N} 1_2^T \Sigma^{-1} 1_2)^{-1}$$

(3) Implement the algorithm in question (2) to obtain $\hat{\mu}_{MLE}$ and its variance $V(\hat{\mu}_{MLE})$. Also obtain $\hat{\Sigma}_{MLE}$.
Solution:

```
> Sigma <- diag(n)
> one <- matrix(1,2,1)
> for(i in 1:10){
+   mu <- as.numeric(solve(N*t(one)%*%solve(Sigma)%*%one)%*%
+       (apply(t(one)%*%solve(Sigma)%*%t(dataset),1,sum)))
+   Sigma <- t(dataset-mu)%*%(dataset-mu)/N
+ }
> mu
[1] 0.01756849

> Sigma
          [,1]      [,2]
[1,] 2.1347913 0.5384672
[2,] 0.5384672 1.0835301

> v_mu <- solve(N*t(one)%*%solve(Sigma)%*%one)
> v_mu
             [,1]
[1,] 0.0009447913
> q <- N*(t(one)*mu)%*%solve(Sigma)%*%one*mu
         [,1]
[1,] 0.326688
> pchisq(q,1,lower.tail=F)
          [,1]
[1,] 0.5663062
```

(4) Is $\hat{\Sigma}_{MLE}$ close to what you expect? Is $\hat{\mu}_{MLE}$ close to what you expect? What is the p-value for testing $H_0 : \mu = 0$?

Solution: Yes, the structure of covariance matrix $\hat{\Sigma}_{MLE}$ should be similar to the covariance matrix used to generate data. Yes, the elements of $\hat{\mu}_{MLE}$ should be all approximate to 0. We use chi-square test to test the hypothesis.

$$N\hat{\mu}^T(\hat{\Sigma})^{-1}\hat{\mu} \sim \chi_1^2$$

The p-value is about 0.5663062.

(5) Now we take a step back. Instead of using the normality assumption, we will use the generalized estimating equation (a.k.a weighted least squares) to perform estimation. Assume that our working covariate matrix is of this form: $\Sigma = \begin{bmatrix} \sigma^2 & \rho \\ \rho & \sigma^2 \end{bmatrix}$, with only two parameters instead of three parameters used in question (2). Write down the generalized estimating equation and an algorithm similar to what we describe in class to estimate (μ, σ^2, ρ).
Solution: We have

$$Q_W(\mu) = \sum_{i=1}^{N}(Y_i - 1_2\mu)^T W_i(Y_i - 1_2\mu)$$

$$\text{Let } \frac{\partial Q_W(\mu)}{\mu} = -2\sum_{i=1}^{N}(1_2^T W_i Y_i - \mu' 1_2^T W_i 1_2) = 0$$

$$\Rightarrow \hat{\mu} = (\sum_{i=1}^{N} 1_2^T W_i 1_2)^{-1} \sum_{i=1}^{N} 1_2^T W_i Y_i$$

Also, the Σ can be estimated by

$$\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - 1_2\mu)(Y_i - 1_2\mu)^T$$

Then, the diagonal elements of Σ can be estimated by averaging the diagonal elements of $\hat{\Sigma}$ and the off-diagonal elements of Σ can be estimated by averaging the off-diagonal elements of $\hat{\Sigma}$.

$$\hat{\sigma}^2 = \frac{1}{2N}\sum_{i=1}^{N}\sum_{j=1}^{2}(Y_{ij} - 1_{2ij}\hat{\mu})^2$$

$$\hat{\rho} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1,l=2}(Y_{ij} - 1_{2ij}\hat{\mu})(Y_{il} - 1_{2il}\hat{\mu}) = \frac{1}{N}\sum_{i=1}^{N}(Y_{i1} - 1_{2i1}\hat{\mu})(Y_{i2} - 1_{2i2}\hat{\mu})$$

The algorithm is summarized as follow:

Algorithm 2 WLS Algorithm

1: Given $Y_1, Y_2, \ldots, Y_{1000}$. Set $\hat{W}_{(0)i} = I$.
2: for t in 1:T:
3: $\hat{\mu}_{(t)} = (\sum_{i=1}^{N} 1_2^T \hat{W}_{(t-1)i} 1_2)^{-1} \sum_{i=1}^{N} 1_2^T \hat{W}_{(t-1)i} Y_i$.
4: $\hat{\Sigma}_{(t)} = \frac{1}{N} \sum_{i=1}^{N} (Y_i - 1_2\mu_{(t)})(Y_i - 1_2\mu_{(t)})^T$.
5: $\hat{\sigma}^2_{(t)} = mean(diag(\hat{\Sigma}_{(t)}))$.
6: $\hat{\rho}_{(t)} = \hat{\Sigma}_{(t)1,2}$.
7: $\hat{\Sigma}_{(t)} = \begin{bmatrix} \sigma^2_{(t)} & \hat{\rho}_{(t)} \\ \hat{\rho}_{(t)} & \sigma^2_{(t)} \end{bmatrix}$.
8: $_{(t)} = (\hat{\Sigma}_{(t)})^{-1}$.
9: Until convergence.

(6) Implement the algorithm in question (5) to obtain $\hat{\mu}_{WLS}$. Obtain the robust variance estimate $V(\hat{\mu}_{WLS})$. Also, obtain the estimate $\hat{\Sigma}_{WLS}$.
Solution:

```
> W <- diag(n)
> one <- matrix(1,2,1)
>
> for(i in 1:10){
+   mu <- as.numeric(solve(N*t(one)%*%
+     W%*%one)%*%apply(t(one)%*%W%*%t(dataset),1,sum))
+   Sigma <- t(dataset-mu)%*%(dataset-mu)/N
+   sigma2 <- mean(diag(Sigma))
+   rho <- Sigma[1,2]
+   Sigma <- matrix(rho,2,2)
+   diag(Sigma) <- sigma2
+   W <- solve(Sigma)
+ }
> mu
[1] 0.0152233
> Sigma
          [,1]      [,2]
[1,] 1.6091552 0.5384617
[2,] 0.5384617 1.6091552
> V_model <- solve(t(one)%*%solve(Sigma)%*%one)/N
> V_robust <- (1/N)*(solve(t(one)%*%solve(Sigma)%*%one))%*%
+   (t(one)%*%solve(Sigma)%*%t(dataset-mu)%*%(dataset-mu)%*%
+   solve(Sigma)%*%one)%*%
+   (solve(t(one)%*%solve(Sigma)%*%one))*(1/N)
> V_model
```

```
             [,1]
[1,] 0.001073808
> V_robust
             [,1]
[1,] 0.001073808
```

(7) Is $\hat{\Sigma}_{WLS}$ close to what you expect? Is $\hat{\mu}_{WLS}$ close to what you expect? Which of the variances are bigger: $V(\hat{\mu}_{MLE})$ or $V(\hat{\mu}_{WLS})$? Does the comparison between $V(\hat{\mu}_{MLE})$ and $V(\hat{\mu}_{WLS})$ fit your expectation?
Solution: $\hat{\Sigma}_{WLS}$ is close to my expectation because the orginal diagonal elements are 2 and 1, and in our estimation, the 1.6 is about the average of diagonal elements. $\hat{\mu}_{WLS}$ is close to what I expect as it is close to 0. $V(\hat{\mu}_{WLS}) = 0.001073808 > 0.0009447913 = V(\hat{\mu}_{MLE})$. It fits to my expectation. First because our model about Σ is wrong in WLS, so the variance will be larger; What's more, the MLE has model assumption (in our case is normal), which is a right assumption because our data was generated by normal!

(8) Now compute the model-based variance estimate $V'(\hat{\mu}_{WLS})$ from the WLS algorithm. Assuming that $V'(\hat{\mu}_{WLS})$ is different from $V(\hat{\mu}_{WLS})$, which variance you use in practice, and why?
Solution: $V'(\hat{\mu}_{WLS}) = 0.001073808$. I will use robust variance because our model about Σ is wrong!

Problem 3: Problem 5.1 on the textbook (page 140-141)

I used both R and SAS in this question for convenience.

1. Read the data from the external file and keep it in a "multivariate" or "wide" format.
Solution:

```
> multivariate <- read.table("cholesterol.txt", na.strings='.')
> colnames(multivariate) <- c("Group", "ID", "Y1",
+   "Y2", "Y3", "Y4", "Y5")
```

2. Calculate the sample means, standard deviations, and variances of the serum cholesterol levels at each occasion for each treatment group.
Solution:

```
> apply(multivariate[multivariate$Group==1,][,-c(1,2)],
+       2, mean, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
```

```
226.0161 245.5323 252.0182 256.7955 254.5526
> apply(multivariate[multivariate$Group==2,][,-c(1,2)],
+       2, mean, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
235.9268 243.1707 244.7632 257.6000 257.4839
> apply(multivariate[multivariate$Group==1,][,-c(1,2)],
+       2, sd, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
39.66437 39.45228 38.32922 34.48935 49.96198
> apply(multivariate[multivariate$Group==2,][,-c(1,2)],
+       2, sd, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
55.87459 49.23967 46.11058 51.14179 49.38817
> apply(multivariate[multivariate$Group==1,][,-c(1,2)],
+       2, var, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
1573.262 1556.483 1469.129 1189.515 2496.200
> apply(multivariate[multivariate$Group==2,][,-c(1,2)],
+       2, var, na.rm=TRUE)
      Y1       Y2       Y3       Y4       Y5
3121.970 2424.545 2126.186 2615.482 2439.191
```

3. One a single graph, construct a time plot that displays the mean serum cholesterol versus time (in months) for the two treatment groups. Describe the general characteristics of the time trends for the two groups.
   Solution:

```
> highDose <- as.vector(apply(multivariate[multivariate$Group==1,][
+   ,-c(1,2)], 2, mean, na.rm=TRUE))
> placebo <- as.vector(apply(multivariate[multivariate$Group==2,][
+   ,-c(1,2)], 2, mean, na.rm=TRUE))
> plot(placebo, type="l", col="blue", ylim=c(225,265))
> lines(highDose, type="l", col="red")
> legend(4.3,235, c("placebo", "treatment"), lty=c(1,1),
+   lwd=c(2.5,2.5), col=c("blue","red"))
```

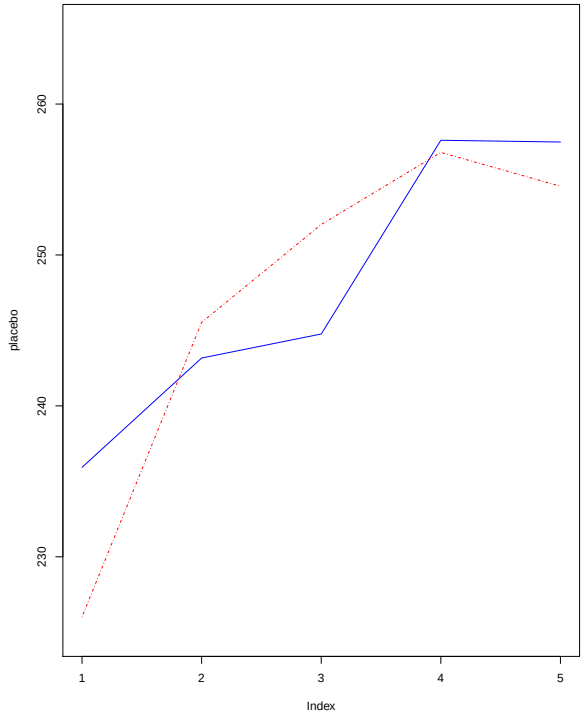

Figure 1: mean serum cholesterol versus time (solid line for placebo)

Treatment group increases fast at first, and gradually becomes slow. Then it decreases. Placebo group is continuously increasing with different rates.

4. Next read the data from the external file and put the data in a "univariate" or "long" format, with five "records" per subject.
Solution:

```
> univariate <- reshape(multivariate,
                        varying = c("Y1", "Y2", "Y3", "Y4", "Y5"),
                        v.names = "level",
                        timevar = "months",
                        idvar=c("ID","Group"),
                        times = c("Y1", "Y2", "Y3", "Y4", "Y5"),
                        direction = "long")

> univariate  <- univariate[order(univariate$ID),]
> rownames(univariate)<-NULL
```

5. Assuming an unstructured covariance matrix, conduct an analysis of response profiles. Determining whether the pattern of change over time differ in the two treatment groups.
Solution: We would like to test $H_0 : \beta_1 - \beta_6 = \beta_2 - \beta_7 = \beta_3 - \beta_8 = \beta_4 - \beta_9 = \beta_5 - \beta_{10}$. The SAS code is as follow

```
libname hw653 'C:\Users\maxyxb\Downloads';

data hw653.cholesterol_long;
set hw653.cholesterol;
highDose = group; /*rename group variable*/
time = 0; level = y1; output;
time = 6; level = y2; output;
time = 12; level = y3; output;
time = 20; level = y4; output;
time = 24; level = y5; output;
drop y1 y2 y3 y4 y5 group;
run;

proc mixed data = hw653.cholesterol_long method=ML;
class highDose(ref="2") time(ref="0") id;
model level = time highDose time*highDose/solution;
repeated/type = un subject = id r rcorr;
run;
```

The result is

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
time	4	101	15.32	<.0001
highDose	1	101	0.05	0.8185
highDose*time	4	101	2.01	0.0990

Figure 2: testing for the pattern of change over time differ in the two groups

Because the interaction term has p value 0.0904, we fail to reject the null hypothesis that the patterns of change over time are the same.

6. Display the estimated 5×5 covariance and correlation matrices for the five repeated measurement of serum cholesterol.
Solution:

```
Estimated R Matrix for ID 1
Row Col1 Col2 Col3 Col4 Col5
1 2144.15 1539.56 1380.17 1416.49 1300.45
2 1539.56 1863.37 1363.52 1426.72 1378.73
3 1380.17 1363.52 1658.63 1225.65 1321.06
4 1416.49 1426.72 1225.65 1718.87 1239.06
5 1300.45 1378.73 1321.06 1239.06 2265.33
```

```
Estimated R Correlation Matrix for ID 1
Row Col1 Col2 Col3 Col4 Col5
1 1.0000 0.7702 0.7319 0.7378 0.5901
2 0.7702 1.0000 0.7756 0.7972 0.6711
3 0.7319 0.7756 1.0000 0.7259 0.6815
4 0.7378 0.7972 0.7259 1.0000 0.6279
5 0.5901 0.6711 0.6815 0.6279 1.0000
```

7. With baseline (month 0) and the placebo group (group 2) as the reference group, write out the regression model for mean serum cholesterol that corresponds to the analysis of response profiles in Problem 5.1.5.
Solution: Our model can be written as

$$\begin{aligned} E(Y_i|X_i) = \beta_0 &+ \beta_1 I(month\ 6) + \beta_2 I(month\ 12) + \beta_3 I(month\ 20) \\ &+ \beta_4 I(month\ 24) + \beta_5 I(group\ 1) + \beta_6 I(group\ 1) I(month\ 6) \\ &+ \beta_7 I(group\ 1) I(month\ 12) + \beta_8 I(group\ 1) I(month\ 20) \\ &+ \beta_9 I(group\ 1) I(month\ 24) \end{aligned}$$

8. Let L denote a matrix of known weights and β the vector of linear regression parameters from the model assumed in Problem 5.1.7. The null hypothesis that the patterns of change over time do not differ in the two treatment groups can be expressed as $H_0 : L\beta = 0$. Describe an appropriate weight matrix L for this null hypothesis.
Solution: Our goal is to test $\beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$, we thus can use contrast matrix

$$L = \begin{bmatrix} 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0 \\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1 \end{bmatrix}$$

9. Show how the estimated regression coefficients from an analysis of response profiles can be used to construct the time-specific means in the two groups. Compare these estimated means with the sample means obtained in Problem 5.1.2.
Solution:

Solution for Fixed Effects

Effect	highDose	time	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			235.93	7.2316	101	32.62	<.0001
time		6	7.2439	4.7586	101	1.52	0.1311
time		12	8.8483	5.1533	101	1.72	0.0890
time		20	23.1028	5.2367	101	4.41	<.0001
time		24	21.1237	7.2738	101	2.90	0.0045
time		0	0	.	.	.	.
highDose	1		-9.9107	9.3209	101	-1.06	0.2902
highDose	2		0	.	.	.	.
highDose*time	1	6	12.2722	6.1334	101	2.00	0.0481
highDose*time	1	12	16.4165	6.6734	101	2.46	0.0156
highDose*time	1	20	4.9765	6.9016	101	0.72	0.4725
highDose*time	1	24	6.9038	9.6626	101	0.71	0.4766
highDose*time	1	0	0	.	.	.	.
highDose*time	2	6	0	.	.	.	.
highDose*time	2	12	0	.	.	.	.
highDose*time	2	20	0	.	.	.	.
highDose*time	2	24	0	.	.	.	.
highDose*time	2	0	0	.	.	.	.

Figure 3: parameters estimation

The sample means obtained in 5.1.2 are as follows.

```
Group 1
      Y1        Y2        Y3        Y4        Y5
226.0161 245.5323 252.0182 256.7955 254.5526
Group 2
      Y1        Y2        Y3        Y4        Y5
235.9268 243.1707 244.7632 257.6000 257.4839
```

We thus can use the estimated parameters construct the time-specific means in the two groups. Take group 2 as example, we have $\hat{\beta}_0 = 235.93, \hat{\beta}_0 + \hat{\beta}_1 = 235.93 + 7.24 = 243.17, \hat{\beta}_0 + \hat{\beta}_2 = 244.76, \hat{\beta}_0 + \hat{\beta}_3 = 257.60, \hat{\beta}_0 + \hat{\beta}_4 = 257.48$, which are similar to the mean value estimated above. Take group as example, we have $\beta_0 - \beta_5 = 226.02, \beta_0 - \beta_5 + \beta_1 + \beta_6 = 245.53, \beta_0 - \beta_5 + \beta_2 + \beta_7 = 252.02, \beta_0 - \beta_5 + \beta_3 + \beta_8 = 256.79, \beta_0 - \beta_5 + \beta_4 + \beta_9 = 254.55$.

10. With baseline (month 0) and the placebo group (group 2) as the reference group, provide an interpretation for each of the estimated regression coefficients in terms of the effect of the treatments on the patterns of change in mean serum cholesterol.
Solution: The interpretations are as follows:

β_0 : the mean serum cholesterol levels at month 0 in group 2 is 235.93 ($p < 0.0001$).
β_1 : difference in mean serum cholesterol levels between month 6 and month 0 in group 2 is 7.24.

β_2 : difference in mean serum cholesterol levels between month 12 and month 0 in group 2 is 8.85.
β_3 : difference in mean serum cholesterol levels between month 20 and month 0 in group 2 is 23.10 ($p < 0.0001$).
β_4 : difference in mean serum cholesterol levels between month 24 and month 0 in group 2 is 21.12 (p=0.0045).
β_5 : difference in mean serum cholesterol levels in group 1 is 9.91 lower than group 2 at month 0.
β_6 : the time adjusted difference at month 6 in mean serum cholesterol levels for group 1 is 12.27 (p=0.0481).
β_7 : the time adjusted difference at month 12 in mean serum cholesterol levels for group 1 is 16.42 (p=0.0156).
β_8 : the time adjusted difference at month 20 in mean serum cholesterol levels for group 1 is 4.98.
β_9 : the time adjusted difference at month 24 in mean serum cholesterol levels for group 1 is 6.90.