# BIOSTAT 653 Homework #2 Solutions

Fall 2017

**Problem 1**

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

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

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

Let $q = c^T (X^T W X)^{-1} X^T W$, $k = c^T (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1}$, and define $r$ such that $q = k + r$.
Note that $kX = qX \Rightarrow rX = \mathbf{0}$.

$$\begin{aligned} V\left(c^T \hat{\beta}_W\right) &= q\Sigma q^T \\ &= (k+r)\Sigma(k+r)^T \\ &= k\Sigma k^T + 2r\Sigma k^T + r\Sigma r^T \\ &= c^T (X^T \Sigma^{-1} X)^{-1} c + 2rX(X^T \Sigma^{-1} X)^{-1} c + r\Sigma r^T \\ &= V\left(c^T \hat{\beta}_{\Sigma^{-1}}\right) + r\Sigma r^T \end{aligned}$$

And since $\Sigma$ is positive definite, $r\Sigma r^T \geq 0$ and $V\left(c^T \hat{\beta}_{\Sigma^{-1}}\right)$ is the smallest possible variance

Note: the hint directly jumps you to the last line, with $r = c^T((X^T W X)^{-1} X^T W - (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1})$. If you skipped directly to the hint, you should go see the beauty of things canceling out when you expand the hint. Also, I bet a bunch of you thought $(A-B)C(A-B)^T = ACA^T - BCB^T$ but this is generally not true.

**Problem 2**

a. Example:

```
library(mvtnorm)
Y=rmvnorm(1000, c(0, 0), matrix(c(2, 0.5, 0.5, 1), 2, 2) )
```

b.

Find the likelihood and take derivatives w.r.t. μ and ∑ and set score to 0, you find:

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

$\hat{\Sigma} = \frac{1}{1000}(Y - 1_{1000\times 2}\mu)^T (Y - 1_{1000\times 2}\mu)$ [Remember Y is 1000 x 2]

Differentiate score w.r.t. μ once more to find information for μ:

$$\hat{V}(\hat{\mu}) = \left(1000(1_2^T \Sigma^{-1} 1_2)\right)^{-1}$$

c.

```
# MLE Algorithm:
Si=diag(2); nidv=1000
for (i in 1:10) {
mu=sum(apply(Si%*%t(Y), 1, sum)/(nidv*sum(Si)))
S=t(Y-mu)%*%(Y-mu)/nidv
Si=solve(S)
}
se_mu=sqrt(1/(sum(Si)*nidv))
```

Hopefully you got something where μ is close to 0 and your ∑ is close to $\begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}$

d.

$\frac{\hat{\mu}^2}{\hat{V}(\hat{\mu})} \sim X_1^2$

Plug in whatever you got.

e.

Similar to part b, but you use a W matrix instead of ∑ to estimate μ, and you average the diagonal of ∑ to find $\sigma^2$. General pseudocode is:

1. Initialize W
2. Estimate μ with W
3. Estimate ∑ with μ
4. Estimate W with ∑
5. Repeat 2-4 until convergence or some iteration limit

```
# WLS Algorithm:
Si=diag(2); nidv=1000
for (i in 1:10) {
mu=sum(apply(Si%*%t(Y), 1, sum)/(nidv*sum(Si)))
S=t(Y-mu)%*%(Y-mu)/nidv
sigma2=mean(diag(S))
rho=S[1,2]
S=matrix(rho, 2, 2)
diag(S)=sigma2
Si=solve(S)
```

}

se_model=sqrt(1/(sum(Si)*nidv))

se_robust=sqrt(sum(Si%*%t(Y-mu)%*%(Y-mu)%*%Si)/(sum(Si)*nidv)^2).

f.

Hopefully you got μ is still close to 0 and your ∑ close to $\begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$.

g.

You should get $\mathrm{V}(\hat{\mu}_{WLS})$ is slightly larger than $\mathrm{V}(\hat{\mu}_{MLE})$, depending on rounding and simulation luck.

h.

Subjective. Do you believe the variance structure? Do you think your sample is large enough? Personally, I'd keep a robust variance with this large of a sample size, but there are other numerical values to compare your different models if you wanted more justification.

**Problem 3**

**Solution**

5.1.1

```
DATA cholesterol;
      INFILE 'cholesterol-data.txt';
      INPUT group id $ Y1 Y2 Y3 Y4 Y5;
RUN;
```

5.1.2

```
PROC MEANS DATA=cholesterol MEAN STD VAR;
CLASS group;

RUN;
```

| group | N Obs | Variable | Mean | Std Dev | Variance |
|---|---|---|---|---|---|
| 1 | 62 | Y1 | 226.0161290 | 39.6643673 | 1573.26 |
| | | Y2 | 245.5322581 | 39.4522819 | 1556.48 |
| | | Y3 | 252.0181818 | 38.3292224 | 1469.13 |
| | | Y4 | 256.7954545 | 34.4893509 | 1189.52 |
| | | Y5 | 254.5526316 | 49.9619841 | 2496.20 |
| 2 | 41 | Y1 | 235.9268293 | 55.8745874 | 3121.97 |
| | | Y2 | 243.1707317 | 49.2396702 | 2424.55 |
| | | Y3 | 244.7631579 | 46.1105805 | 2126.19 |
| | | Y4 | 257.6000000 | 51.1417868 | 2615.48 |
| | | Y5 | 257.4838710 | 49.3881706 | 2439.19 |

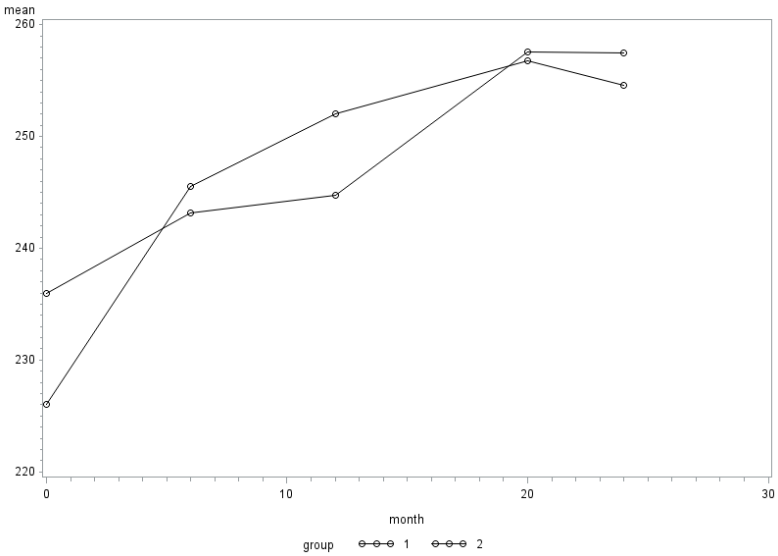
5.1.3 Group 1 increases fast at the beginning, slows down gradually and eventually decreases. Group 2 has a linear increasing trend.

```
DATA cholesterol_long;
      SET cholesterol;
      month=0;
      Y=Y1;
      output;
      month=6;
      Y=Y2;
      output;
      month=12;
      Y=Y3;
      output;
      month=20;
      Y=Y4;
      output;
      month=24;
      Y=Y5;
      output;
      DROP Y1 Y2 Y3 Y4 Y5;
RUN;

PROC SORT DATA=cholesterol_long;
      BY group month;
RUN;

PROC MEANS DATA=cholesterol_long NOPRINT;
      BY group month;
      VAR Y;
      OUTPUT OUT=cholesterol_mean mean=mean;
RUN;
```

```
PROC GPLOT DATA=cholesterol_mean;
      PLOT mean*month=group;
RUN;
```




5.1.4 See the code in 5.1.3.

5.1.5 Because for the interaction term p=0.0904, we fail to reject the null hypothesis that the two patterns of change are the same. Notice that you may get a slight different results if you use REML.

```
PROC MIXED DATA=cholesterol_long METHOD=ML;
      CLASS group month id;
      MODEL Y=month group month*group /S CHISQ;
      REPEATED month/TYPE=un SUBJECT=id R RCORR;
RUN;
```

**Type 3 Tests of Fixed Effects**

| Effect | Num DF | Den DF | Chi-Square | F Value | Pr > ChiSq | Pr > F |
|---|---|---|---|---|---|---|
| **month** | 4 | 101 | 61.28 | 15.32 | <.0001 | <.0001 |
| **group** | 1 | 101 | 0.05 | 0.05 | 0.8181 | 0.8185 |
| **group*month** | 4 | 101 | 8.03 | 2.01 | 0.0904 | 0.0990 |

5.1.6

**Estimated R Matrix for id 1**

| Row | Col1 | Col2 | Col3 | Col4 | Col5 |
|---|---|---|---|---|---|
| **1** | 2144.15 | 1539.56 | 1380.17 | 1416.49 | 1300.45 |

**Estimated R Matrix for id 1**

| Row | Col1 | Col2 | Col3 | Col4 | Col5 |
|---|---|---|---|---|---|
| 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 |

5.1.7 Before, we have

$$Y_{ij} = \beta_0 I(g=0)I(m=0) + \beta_1 I(g=0)I(m=6) + \beta_2 I(g=0)I(m=12) + \beta_3 I(g=0)I(m=20) + \beta_4 I(g=0)I(m=24) + \beta_5 I(g=1)I(m=0) + \beta_6 I(g=1)I(m=6) + \beta_7 I(g=1)I(m=12) + \beta_8 I(g=1)I(m=20) + \beta_9 I(g=1)I(m=24)$$

Now, we have

$$Y_{ij} = \beta_0 + \beta_1 I(m=6) + \beta_2 I(m=12) + \beta_3 I(m=20) + \beta_4 I(m=24) + \beta_5 I(g=1)I(m=0) + \beta_6 I(g=1)I(m=6) + \beta_7 I(g=1)I(m=12) + \beta_8 I(g=1)I(m=20) + \beta_9 I(g=1)I(m=24)$$

5.1.8 To test $\beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$, we use the following contrast matrix L

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

5.1.9 The estimates are for the time-specific means are

| group | Variable | Mean |
|---|---|---|
| 1 | Y1 | 226.0156 |
| | Y2 | 245.5318 |
| | Y3 | 251.2805 |
| | Y4 | 254.0949 |

| group | Variable | Mean |
|---|---|---|
| | **Y5** | 254.0431 |
| **2** | **Y1** | 235.9263 |
| | **Y2** | 243.1702 |
| | **Y3** | 244.7746 |
| | **Y4** | 259.0291 |
| | **Y5** | 257.05 |

which are computed based on the following table:

**Solution for Fixed Effects**

| Effect | group | month | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|---|
| **Intercept** | | | 257.05 | 8.0030 | 101 | 32.12 | <.0001 |
| **month** | | **0** | -21.1237 | 7.2738 | 101 | -2.90 | 0.0045 |
| **month** | | **6** | -13.8798 | 6.4993 | 101 | -2.14 | 0.0351 |
| **month** | | **12** | -12.2754 | 6.3380 | 101 | -1.94 | 0.0556 |
| **month** | | **20** | 1.9791 | 6.8479 | 101 | 0.29 | 0.7732 |
| **month** | | **24** | 0 | . | . | . | . |
| **group** | **1** | | -3.0069 | 10.5770 | 101 | -0.28 | 0.7768 |
| **group** | **2** | | 0 | . | . | . | . |
| **group*month** | **1** | **0** | -6.9038 | 9.6626 | 101 | -0.71 | 0.4766 |
| **group*month** | **1** | **6** | 5.3685 | 8.6974 | 101 | 0.62 | 0.5385 |
| **group*month** | **1** | **12** | 9.5128 | 8.5260 | 101 | 1.12 | 0.2672 |
| **group*month** | **1** | **20** | -1.9273 | 9.2123 | 101 | -0.21 | 0.8347 |
| **group*month** | **1** | **24** | 0 | . | . | . | . |
| **group*month** | **2** | **0** | 0 | . | . | . | . |
| **group*month** | **2** | **6** | 0 | . | . | . | . |
| **group*month** | **2** | **12** | 0 | . | . | . | . |
| **group*month** | **2** | **20** | 0 | . | . | . | . |
| **group*month** | **2** | **24** | 0 | . | . | . | |

The values are almost identical to the sample means. The small difference is due to the fact that we obtained the mean estimates via a model based approach here, and our model accounts for correlation among repeated measurements.

5.1.10

Interpretations are generally very flexible depending on what model you chose, but whatever you say as your interpretation, you should make sure the numbers match up. Don't say that time 6 group 1 and time 0 group 2 differ by 5 when 245 and 235 do not differ by 5.
Make sure you know when you're using cell means and when you're not.