ST 437/537 Homework 5, Spring 2018

ST 537: Q1-10. ST 437: Q1-6

[*Background*]. A common way of treating patients with cardiovascular disease is by surgical intervention. In particular, such patients may arrive at a hospital with symptoms such as unstable angina or suspected myocardial infarction (heart attack), requiring that physicians perform an invasive procedure called a percutaneous coronary intervention (PCI) to investigate the extent to which coronary arteries might be blocked. During such an investigation, the blockage may be treated using a balloon to dislodge the blockage and widen the artery ("balloon angioplasty"); in addition, a device known as a stent may be inserted to prop the artery open.

When such PCI procedures are performed, it is necessary for the subject to be treated with a drug that inhibits the aggregation of platelets in the blood. Informally, platelets are a blood constituent involved in clotting of the blood; clotting occurs when the platelets aggregate together in "clumps." To ensure that clotting does not interfere with the procedure, inhibition of the clotting mechanism is desirable; clotting during the procedure can lead to complications such as stroke or heart attack. A long-standing issue has been to determine which of two popular drugs elicits the most desirable pattern of inhibition of platelet aggregation.

[*Experiment*]. Accordingly, an experiment was conducted to compare the platelet aggregation patterns of the drugs in such subjects under controlled conditions. Subjects arriving at a major medical center with symptoms of unstable angina or myocardial infarction who were judged to require a PCI procedure were randomized into two groups, one for each of the drugs, with 200 subjects per group. For each subject, at time 0, the assigned drug was administered according to the manufacturer's recommended dosage; for each drug, this involved giving the subject a large dose by injection to start inhibition of platelet aggregation immediately and simultaneously giving the subject a smaller dose of the drug intravenously at a constant rate over several hours, a method of administration known as an infusion. The purpose of the infusion was to keep platelet aggregation inhibited over at least a 12 hour period, so that clotting would be minimized during the PCI procedure and subsequent recovery for the subject.

For each subject, blood samples were to be taken at 0.5, 2.0, 3.5, 5.0, 8.0, 11.0, and 12.5 hours. Each sample was to be analyzed for degree of platelet inhibition, characterized by the response "percent inhibition," a value between 0 and 100 representing the percentage of inhibition relative to that of an untreated sample (in units of "%μM"). Also recorded for each subject was whether the subject had experienced a previous myocardial infarction before the current hospitalization (0=no, 1=yes) and their gender (0=female, 1=male).

[*Data*]. The data from the study is given by `platelet.txt`, found in the "Data" folder located on the coure webpage. Each record in the file corresponds to a single observation, and the columns are (1) subject id number ($i = 1$ to $i = 400$), (2) previous myocardial infarction indicator (m_i, 0 = no, 1 = yes), (3) gender indicator (g_i, 0 = female, 1 = male), (4) time (hours, measured since administration of drug), (5) percent inhibition, and (6) drug group indicator ($k = 1$ or $k = 2$). Note that for some subjects, the response is not available at all intended time points; some samples were mishandled and in some instances study personnel did not follow the instructions and neglected to obtain samples. It was determined that the reasons for the missing values had nothing to do with the drugs or the patterns of inhibition.

1. Plot the data (using appropriate graphics) separately for each treatment drug and comment on the mean inhibition trajectory in each group. Based on this plot, what can you infer about how the mean inhibition changes over time ?

[*Model*]. The investigators were particularly interested in the time point 0.5 hours post-administration; thus we define a "new time" variable, t, so that $t = 0$ corresponds to 0.5 hours after administration of the drug and $t = 12$ corresponds to 12.5 hours after the administration of the drug.

Let Y_{ij} be the corresponding platelet inhibition response for subject i at the jth time.

2. The investigators wished to assume a model for the mean platelet inhibition within each drug group that changes linearly with the new time defined above. That is

$$Y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} + e_{ij}, \tag{1}$$

where β_{0i} is the mean platelet inhibition at time 0.5 hrs following drug administration, β_{1i} is the slope of the rate of change in the subject mean inhibition trajectory, and e_{ij} is the mean zero deviation associated wiht the jth response; e_{ij} are assumed normally distributed.

Furthermore the investigators wished to assume that the mean platelet inhibition has the following structures:

1. is associated with whether the subject has had a previous myocardial infarction
2. is associated with whether the subject is male or female
3. the way in which it is associated with whether the subject is male or female is the different depending on whether the subject has had a previous myocardial infarction.

Because the subjects had been on the drugs for 0.5 hours, the investigators assumed that mean platelet inhibition at 0.5 hours and the way the above features occur is *different* for the two drugs. Also they wished to assume that the "typical" or mean rate of change of platelet inhibition over the study period *also* has these features described above.

Given these beliefs, write down the expressions of β_{0i} and β_{1i} for subject i taking drug k, $k = 1, 2$. Be sure to define all the additional symbols you use.

Hint: for β_{0i} and for each k there will need to be 4 fixed effect β's multiplied appropriately by m_i and g_i to be added to the random b_{0i}. And similarly 4 for each k for the slope, together resulting in a vector $\boldsymbol{\beta}$ of length 16 like

$$\boldsymbol{\beta} = (\beta_{10}, \beta_{20}, \beta_{10g}, \beta_{20g}, \beta_{10m}, \beta_{20m}, \beta_{10gm}, \beta_{20gm}, \beta_{11}, \beta_{21}, \beta_{11g}, \beta_{21g}, \beta_{11m}, \beta_{21m}, \beta_{11gm}, \beta_{21gm})'.$$

3. Using the model described in 2. give an expression that represents the typical value of platelet inhibition at 0.5 hours after drug administration for male subjects with a previous myocardial infarction taking drug 2.

4. Using the model described in 2. give an expression for mean platelet inhibition for female subjects with no previous myocardial infarction at 12 hours following administration of drug 1.

In the following parts, you will write R code to carry out several different analyses of this data. You will have to modify your program for each part to obtain desired analyses. Please turn in your final program and output that carries out all necessary analyses.

The investigators were willing to believe that

(i) the assay used to measure platelet inhibition for both drug groups exhibits constant variation regardless of the true value of platelet inhibition being ascertained,

(ii) within-subject local "fluctuations" in platelet inhibition are of similar magnitude for both drugs and across time for all subjects,

(iii) variation in "inherent," true platelet inhibition at 0.5 hours is similar for patients in both drug groups, as is variation in the "inherent" rates of change of platelet inhibition over the study period and the way these quantities co-vary.

One of the investigators was concerned, however, that the time points at which platelet inhibition was measured were not sufficiently far apart in time to ensure that measurements within a subject are uncorrelated. He was willing to believe that, if such correlation is present, it "falls off" as the time points get farther apart, but he insisted that an analysis be done to resolve this issue.

5. Give two different sets of assumptions on the residuals e_{ij}, $i = 1, \ldots, n_i$ and the random effects corresponding to β_{0i} and β_{1i} in in (1) that incorporate (i)–(iii). The first set of assumptions should incorporate the investigator's concern; the second set should represent the case where the investigator's concern is unwarranted.

   Fit the overall model (1) along with your model for β_{0i} and β_{1i} in (2) and under both sets of assumptions using R `nlme package`. Which set of assumptions is best supported?

6. From the output for the fit of the model you preferred in 5., write down an estimate of the variance associated with among-subject variation in true platelet inhibition in the population of male subjects with no previous myocardial infarction receiving drug 2 at 0.5 hours post-administration.

7. Previous research has suggested that the way in which platelet inhibition occurs for both drugs over this period may be associated with whether a subject has had a previous myocardial infarction, but there is no evidence to suggest that it is associated with gender in any way. Thus, the investigators planned to base their subsequent analyses not on the mean model you developed in 2. but on a mean model that includes no effect of gender either in the representation of mean platelet inhibition at 0.5 hours or in the representation of the "typical" rate of change of platelet inhibition over the study period. Write down this simpler model and fit it using ML and your preferred covariance structure from 5. Based on your preferred fit in 5. and this fit, is there any evidence against doing this? (If you used REML in 5., just rerun it here with `method=ml`.)

8. For the rest of the problem, consider the simpler model in 7. with no gender effects. The reason that the investigators were so interested in 0.5 hours post-administration is because another research team had recently published a paper receiving a lot of press, which claimed that the 2 drugs exhibit the same mean platelet inhibition and that, furthermore, mean platelet inhibition on the two drugs is the same for subjects with or without a previous myocardial infarction. This team based their finding on comparing platelet inhibition levels 0.5 hours post-administration. Our investigators felt that comparing platelet inhibition at a single time point, particularly one so soon after administration, was not very informative.

However, their first goal was to examine whether the data from the current study offer evidence refuting the claim of their rival investigators.

Write down a set of hypotheses that addresses the issue of interest to the investigators in terms of the model in 7., and express your null hypothesis in terms of a linear function $\boldsymbol{L\beta}$, defining $\boldsymbol{L}$. Using Wald methods, carry out the test at level of significance 0.05 based on a REML fit. State your conclusion as a *meaningful sentence*.

9. The investigators' second goal was to make the point that comparing platelet inhibition at a single point does not tell the whole story. Thus, regardless of how the test in 8. turned out, they wanted to investigate longer time periods and the rate of change of platelet inhibition over them. The first question along these lines was whether the way "typical" rate of change differs between subjects who have had a previous myocardial infarction and those who have not, is different for the two drugs. (This is an interaction!)

   Write down a set of hypotheses that addresses the issue of interest to the investigators in terms of the model in 7., and express your null hypothesis in terms of a linear function $\boldsymbol{L\beta}$, defining $\boldsymbol{L}$. Using Wald methods, carry out the test at level of significance 0.05 based on a REML fit. State your conclusion as a *meaningful sentence*.

10. Based on the model in 7., provide estimates (and associated standard errors) of mean platelet inhibition at 12.5 hours after administration of (a) drug 1 in subjects with previous myocardial infarction; and (b) drug 2 in subjects with no previous myocardial infarction.