CPH 931 — Fall 2008 — Dr. Charnigo

Written Assignment 2 Solutions

1a. The revised linear mixed model is given by

$$Y_{ij} = \beta_0 + \sum_{k=1}^{m} \beta_k x_{k,ij} + \alpha_i + \epsilon_{ij}$$

with $m = 12$ and $X_1, \ldots, X_{12}$ defined as follows: $X_1 = 1$ when DRUG = 0, $X_2 = 1$ when ACTIVITY = 0, $X_3 = 1$ when TIME = 2, $X_4 = 1$ when TIME = 3, $X_5 = 1$ when SSSTAT = 0, $X_6 = 1$ when DRUG = 0 and SSSTAT = 0, $X_7 = 1$ when ACTIVITY = 0 and SSSTAT = 0, $X_8 = 1$ when SSSTAT = 0 and TIME = 2, $X_9 = 1$ when SSSTAT = 0 and TIME = 3, X_{10} = CORRECTTRIALS_B, $X_{11} = 1$ when DRUG = 0 and ACTIVITY = 0, $X_{12} = 1$ when DRUG = 0 and SSSTAT = 0 and ACTIVITY = 0.

1b. The null hypothesis that DRUG and ACTIVITY do not interact within either stratum of SSSTAT is $\beta_{11} = \beta_{12} = 0$. The result, using obvious notation, is $f(2, 357) = 2.32, p = 0.0998$. We cannot reject the null hypothesis.

The SAS code I used to perform the hypothesis test is available in the code file {WA2931F08SAS.txt}: interaction within SSSTAT = 0 is quantified as the difference between the drug effect at high activity when SSSTAT = 0 ("d - b" in the notation of the code file) and the drug effect at low activity when SSSTAT = 0 ("c - a"), while interaction within SSSTAT = 1 is quantified as the difference between the drug effect at high activity when SSSTAT = 1 ("h - f") and the drug effect at low activity when SSSTAT = 1 ("g - e").

1c. The null hypothesis that ACTIVITY is altogether irrelevant is $\beta_2 = \beta_7 = \beta_{11} = \beta_{12} = 0$. The result is $f(4, 357) = 2.20$, $p = 0.0684$. We cannot reject the null hypothesis.

The SAS code I used to perform the hypothesis test is available in the code file {WA2931F08SAS.txt}: the activity effect when SSSTAT = 0 and DRUG = 0 is the expected difference between the high activity score and the low activity score when SSSTAT = 0 and DRUG = 0 ("b - a"), the activity effect when SSSTAT = 0 and DRUG = 1 is the expected difference between the high activity score and the low activity score when SSSTAT = 0 and DRUG = 1 ("d - c"), the activity effect when SSSTAT = 1 and DRUG = 0 is the expected difference between the high activity score and the low activity score when SSSTAT = 1 and DRUG = 0 ("f - e"), and the activity effect when SSSTAT = 1 and DRUG = 1 is the expected difference between the high activity score and the low activity score when SSSTAT = 1 and DRUG = 1 ("h - g").

Remark: An alternative solution entails performing a likelihood ratio test. Fit the model without ACTIVITY, SSSTAT*ACTIVITY, DRUG*ACTIVITY, and SSSTAT*DRUG*ACTIVITY. Observe that negative twice the log likelihood increases from 2157.0 to 2165.7. The difference of 8.7 is less than the $9.49 = \chi^2_{4,0.95}$ that would be required for rejection of the null hypothesis.

1d. Using the SAS code in {WA2931F08SAS.txt}, I estimated the expected difference as -2.6663 with corresponding p-value 0.0098, so the estimate is significantly different from 0.

1e. Using the SAS code in {WA2931F08SAS.txt}, I estimated the expected difference as -0.2449 with corresponding p-value 0.8114, so the estimate is not significantly different from 0.

2a. The parameter estimates and their standard errors are as follows.

Parameter	Estimate	Standard Error
α	1.41622	0.51918
β_1	-0.32046	0.09347
β_2	-0.05253	0.08466

2b. If you took the SEED number to be 12345 when invoking PROC MI, the five filled-in values for Lake Annie were -0.78271, -0.21319, -0.36071, 0.14529, and -0.48495.

2c. If you take the SEED number to be 12345 when invoking PROC MI, the overall parameter estimates and standard errors are as follows. They are very similar to the results reported for part a.

Parameter	Estimate	Standard Error
α	1.638970	0.513141
β_1	-0.358408	0.096199
β_2	-0.038923	0.082810

2d. If you took the SEED number to be 12345 when invoking PROC MI, the five filled-in values for Lake Annie were 0.24942, -0.02157, 0.33936, 0.32127, and 0.23373.

2e. If you take the SEED number to be 12345 when invoking PROC MI, the overall parameter estimates and standard errors are as follows. They are very similar to the results reported for parts a and c, with one exception. The estimate of β_2 is much larger in absolute value, so much so that X_2 is now identified as a significant predictor ($p = 0.0303$).

Parameter	Estimate	Standard Error
α	1.514807	0.474682
β_1	-0.300735	0.087141
β_2	-0.186680	0.085837