

Please read the instructions carefully for this test. You are expected to work on your own on this test and similarity across tests will be closely scrutinized. This test is due at the beginning of our class time (9:00AM) on Wednesday March 17<sup>th</sup>, 2021 and should be e-mailed to me with BST630 Midterm on the subject line. Please be sure to include your name in the file title. On calculation problems, it is always best to show your intermediate steps of calculations. If you do not show these steps, and only provide a final number, I cannot give partial credit for having the correct idea, but making a simple math error.

For the first 3 questions, we will use an example of predicting CD4 cell count over time. We use the square root of CD4 to make the dependent variable more closely to normally distributed. There are 4 time points and two treatment conditions in this data example:

1. When we use only Time as a predictor we get the following results:

**MODEL 1:**

**Covariance Parameter Estimates**

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
TOEP(2)	ID	43.7894	5.5766	7.85	<.0001
TOEP(3)	ID	40.1186	5.5601	7.22	<.0001
TOEP(4)	ID	35.7987	5.6303	6.36	<.0001
Residual		49.8067	5.5909	8.91	<.0001

**Solution for Fixed Effects**

Effect	stime	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept		21.2741	0.6754	124	31.50	<.0001
stime	0	0.01819	0.5338	274	0.03	0.9728
stime	1	0.2508	0.4619	274	0.54	0.5877
stime	2	-0.5901	0.3729	274	-1.58	0.1147
stime	3	0	.	.	.	.

**Covariance Matrix for Fixed Effects**

Row	Effect	stime	Col1	Col2	Col3	Col4	Col5
1	Intercept		0.4561	-0.1700	-0.1248	-0.08026	
2	stime	0	-0.1700	0.2850	0.1893	0.1146	

### Covariance Matrix for Fixed Effects

Row	Effect	stime	Col1	Col2	Col3	Col4	Col5
3	stime	1	-0.1248	0.1893	0.2134	0.1100	
4	stime	2	-0.08026	0.1146	0.1100	0.1391	
5	stime	3					

### Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
stime	3	274	2.20	0.0882

- 1.) How is time being modeled here (5-pts)?
- 2.) Using the results above can you show me what the (residual) covariance matrix over time looks like (i.e. what are the values in the matrix) (5-pts)?
- 3.) What does this model tell you about change over time (use the appropriate statistical tests to back up what you say)(5-pts)?
- 4.) If I reran this model with the following contrast, what am I testing in this contrast (5 pts) (show null and alternative hypothesis)?

contrast "What is it?" stime 0 1 -1 0;

- 5.) Please test the hypothesis in 4 (10 pts)

The next several questions refer to the next set of output:

### MODEL 2:

#### Fit Statistics

-2 Log Likelihood	2297.9
AIC (Smaller is Better)	2321.9
AICC (Smaller is Better)	2322.7
BIC (Smaller is Better)	2356.0

### Solution for Fixed Effects

Effect	RAN group assignment	stime	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept			20.7115	0.9061	123	22.86	<.0001
stime	0		1.1121	0.6974	271	1.59	0.1120

### Solution for Fixed Effects

Effect	RAN group assignment	stime	Estimate	Standard Error	DF	t Value	Pr >  t
stime		1	0.8944	0.6016	271	1.49	0.1383
stime		2	-0.3337	0.4906	271	-0.68	0.4970
stime		3	0	.	.	.	.
TREAT	1		1.3488	1.3515	123	1.00	0.3203
TREAT	2		0	.	.	.	.
TREAT *stime	1	0	-2.4829	1.0496	271	-2.37	0.0187
TREAT *stime	1	1	-1.4643	0.9176	271	-1.60	0.1117
TREAT *stime	1	2	-0.5127	0.7485	271	-0.68	0.4940
TREAT *stime	1	3	0	.	.	.	.
TREAT *stime	2	0	0	.	.	.	.
TREAT *stime	2	1	0	.	.	.	.
TREAT *stime	2	2	0	.	.	.	.
TREAT *stime	2	3	0	.	.	.	.

- 6.) Calculate the means of the two treatment groups at each time (10-pts)
- 7.) Write out a contrast that will test if the difference between conditions at baseline is different from the difference between conditions at time=2 (5-pts).
- 8.) What is the difference in the difference between conditions at time=0 and time=2 (5-pts)?

The next set of questions are about the following model run on the data:

### MODEL 3:

#### Fit Statistics

-2 Log Likelihood	2301.0
AIC (Smaller is Better)	2317.0
AICC (Smaller is Better)	2317.4
BIC (Smaller is Better)	2339.8

### Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr >  Z
UN(1,1)	ID	40.1113	5.6126	7.15	<.0001
UN(2,1)	ID	-0.7858	1.1669	-0.67	0.5007
UN(2,2)	ID	2.0215	0.4804	4.21	<.0001
Residual		5.1749	0.5573	9.29	<.0001

### Solution for Fixed Effects

Effect	RAN group assignment	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept		21.8203	0.8132	123	26.83	<.0001
stime		-0.4858	0.2358	105	-2.06	0.0419
RAN002B	1	-1.0553	1.1939	170	-0.88	0.3780
RAN002B	2	0	.	.	.	.
stime*RAN002B	1	0.8861	0.3584	170	2.47	0.0144
stime*RAN002B	2	0	.	.	.	.

### Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
stime	1	105	0.06	0.8119
RAN002B	1	170	0.78	0.3780
stime*RAN002B	1	170	6.11	0.0144

- 9.) How is time being treated in this model (5-pts)?
- 10.) What do the covariance parameter estimates in this model tell you (10-pts)?
- 11.) Please provide the empirical estimates of the residual variance covariance over time (in the Ys) (10-pts).
- 12.) What do the Fixed Effects estimates tell you about what happened in the study (10-pts)?
- 13.) Which of the two models (Model 2 or Model 3) are better fit to the data and why (5-pts)?

The final question does not refer to any of the prior models:

- 14.) I have a study where I have measured multiple family members (who all live in the same household) in a cross-sectional study. These family members are further nested within

neighborhoods. My primary outcome is psychological distress (DIS). My predictors are Gender (F=1 if female), person income (PI) and Social Support (SS) at the individual level, at the family level we measure family cohesion (FC), and at the neighborhood level we measure amount of violent crime (VC) in the neighborhood. Please put together a statistical model which will account for the following things: 1) family members are not independent, 2) provide a test of whether the individual level variables of gender, person income and social support affect DIS, and 3) whether the Family level variable, family cohesion, and the neighborhood level variable of violent crime affect psychological distress, and 3 whether there is a contextual effect of family income on each individual's psychological distress (DIS). Please write out a theoretical model that would allow me to accomplish all of these things. You may find using the multi-level approach to growth curves helpful in setting this up. Also please briefly explain each component of the model and any assumptions you are making and why. Be sure to tell me which statistical tests will tell me whether each of my conjectures are consistent with the data. Finally, please write out the SAS statements I would use to estimate the model and test these hypotheses. 20 Points

Extra Credit (If you do not need these points for the midterm they can/will be applied to your homework):

How would you modify your model in question 14 to test whether neighborhood violence moderates (interacts with) the effect of person-income on psychological distress?