Grad Stats Exam 2 Spring, 1997

Question 1 (8 points total)

A resarcher is interested in the effects of practice in learning to do a difficult memorization task. Three subjects come in on day one and are given 5 minutes to memorize in order as many digits as they can from a list of 100 digits. After a filler task, they are asked to recite the list in order; the researcher records the number of digits correctly recalled in order before the first error. Subjects return for the same thing on day two and on day three, with a different list of 100 digits on each day. The researcher wants to know if subjects improve over time, even though the lists are different ones.

The data from the three subjects are given below:

	Day 1	Day 2	Day 3
Subject 1	27	29	36
Subject 2	16	18	17
Subject 3	31	30	39

Provide the F* and Pre to test the researcher's hypothesis. (You don't need to answer all questions that could be asked of these data, nor provide a complete source table. Just give us the F* and Pre to ask whether "subjects improve over time.")

What substantive conclusion do you reach?

What advice would you give the researcher, given the size of the Pre in this study?

Question 2 (21 points total, 3 each)

(example taken from Agresti, A., & Finlay, B. (1997). Statistical methods for the social sciences, 3rd edition. NJ: Prentice Hall. p. 525-526)

A researcher is interested in factors associated with fertility in a Latin American city. Of particular interest is whether migrants from other cities or migrants from rural areas differ from natives of the city in their completed family sizes. The groups to be compared are urban natives, urban migrants, and rural migrants. Additionally, the researcher wants to examine how years of education relates to fertility in the three groups.

The variables that are gathered are:

GROUP	RM	=	rural migrants to the city
	UM	=	urban migrants to the city
	UN	=	urban natives of the city
FERT	Complete family size		
EDUC	Years of education completed		

Additionally, the researcher creates the following variables for purposes of analysis:

X1	=	2	if GROUP = UN
	=	-1	otherwise
X2	=	1	if GROUP = UM
	=	-1	if GROUP = RM
	=	0	if GROUP = UN
EDX1	=	X1 * EDUC	
EDX2	=	X2 * EDUC	

On the following pages are:

group means of FERT and EDUC;
simple regressions of FERT on EDUC within each group;
the ANOVA model (regression of FERT on X1 and X2);
a simple regression of FERT on EDUC for the full sample;
the ANCOVA model (regression of FERT on EDUC, X1, and X2);
the ANCOVA model including the interactions.

Based on this output, answer the following questions.

A. Are there differences in fertility among the three groups? (Using the simplest analysis you can, provide the omnibus F* and PRE for the test of group differences.) Based on the contrast codes that are used in the analysis, describe any significant contrasts among the three groups.

B. Overall, is there a significant relationship between educational level and fertility? (Provide F* and PRE and interpret the relevant parameter estimate.)

C. Within the three groups, on average, (or controlling for group membership, is there a significant relationship between educations level and fertility? (Provide F* and PRE and interpret the relevant parameter estimate.)

D. Are there differences in fertility among the three groups controlling for education? (Provide the omnibus F* and PRE for this test). Based on the contrast codes used in the

analysis, describe any signficant contrasts among the three groups, controlling for education.

E. What are the values of the adjusted FERT means for the three groups, controlling for education?

F. In the simple regressions of FERT on EDUC in each group, the slope for EDUC in the UM group is -.12 while for the RM group it is -.38. Are these two slopes significantly different from each other? (Provide F* and PRE and interpret the difference in slopes if they are different.)

G. In the final model that includes the two interactions, provide interpretations for the following three slopes:

-.25	the parameter estimate for EDUC
-1.39	the parameter estimate for X2
.13	the parameter estimate for EDX2

group means

---------- GROUP=rm ----------

Variable	N	Mean	Std Dev	Minimum	Maximum
EDUC	20	4.8000000	3.4580189	0	10.0000000
FERT	20	6.0000000	1.7770466	3.0000000	10.0000000

---------- GROUP=um ----------

Variable	N	Mean	Std Dev	Minimum	Maximum
EDUC	17	6.0588235	4.7889272	0	12.0000000
FERT	17	4.2941176	1.6110829	2.0000000	7.0000000

---------- GROUP=un ----------

Variable	N	Mean	Std Dev	Minimum	Maximum
EDUC	15	5.7333333	4.1484363	0	12.0000000
FERT	15	4.1333333	1.4573296	2.0000000	7.0000000

separate group regression

---------- GROUP=rm ----------

Model: MODEL1
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	33.31426	33.31426	22.471	0.0002
Error	18	26.68574	1.48254		
C Total	19	60.00000			

Root MSE	1.21760	R-square	0.5552
Dep Mean	6.00000	Adj R-sq	0.5305
C.V.	20.29327		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	7.838028	0.47378186	16.544	0.0001
EDUC	1	-0.382923	0.08077914	-4.740	0.0002

separate group regression

---------- GROUP=um ----------

Model: MODEL1
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	5.84057	5.84057	2.455	0.1380
Error	15	35.68884	2.37926		
C Total	16	41.52941			

Root MSE	1.54248	R-square	0.1406
Dep Mean	4.29412	Adj R-sq	0.0833
C.V.	35.92085		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	5.058512	0.61480125	8.228	0.0001
EDUC	1	-0.126162	0.08052345	-1.567	0.1380

separate group regression

---------- GROUP=un ----------

Model: MODEL1
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	14.67744	14.67744	12.673	0.0035
Error	13	15.05589	1.15815		
C Total	14	29.73333			

Root MSE	1.07617	R-square	0.4936
Dep Mean	4.13333	Adj R-sq	0.4547
C.V.	26.03641		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	5.548423	0.48499309	11.440	0.0001
EDUC	1	-0.246818	0.06933191	-3.560	0.0035

combined analyses

Model: ANOVA
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	39.25649	19.62824	7.327	0.0016
Error	49	131.26275	2.67883		
C Total	51	170.51923			

Root MSE	1.63671	R-square	0.2302
Dep Mean	4.90385	Adj R-sq	0.1988
C.V.	33.37612		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	4.809150	0.22854824	21.042	0.0001
X1	1	-0.337908	0.16715553	-2.022	0.0487
X2	1	-0.852941	0.26996292	-3.159	0.0027

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	1186.113791	.	.
X1	1	10.947177	0.07697899	0.99809470
X2	1	26.740859	0.16924208	0.99809470

combined analyses

Model: SIMPLE REG
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	55.12921	55.12921	23.888	0.0001
Error	50	115.39003	2.30780		
C Total	51	170.51923			

Root MSE	1.51914	R-square	0.3233
Dep Mean	4.90385	Adj R-sq	0.3098
C.V.	30.97864		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	6.298841	0.35474497	17.756	0.0001
EDUC	1	-0.254525	0.05207620	-4.888	0.0001

combined analyses

Model: ANCOVA
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	83.75158	27.91719	15.444	0.0001
Error	48	86.76765	1.80766		
C Total	51	170.51923			

Root MSE	1.34449	R-square	0.4912
Dep Mean	4.90385	Adj R-sq	0.4594
C.V.	27.41710		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	6.085810	0.31853117	19.106	0.0001
EDUC	1	-0.230831	0.04652600	-4.961	0.0001
X1	1	-0.314524	0.13739220	-2.289	0.0265
X2	1	-0.707654	0.22368845	-3.164	0.0027

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	659.856427	.	.
EDUC	1	44.495095	0.33897733	0.98130832
X1	1	9.473254	0.09843272	0.99692002
X2	1	18.091351	0.17253026	0.98098881

combined analyses

Model: ANCOVA with Interactions
Dependent Variable: FERT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	93.08876	18.61775	11.060	0.0001
Error	46	77.43048	1.68327		
C Total	51	170.51923			

Root MSE	1.29741	R-square	0.5459
Dep Mean	4.90385	Adj R-sq	0.4966
C.V.	26.45698		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	6.148321	0.30986470	19.842	0.0001
EDUC	1	-0.251968	0.04592570	-5.486	0.0001
X1	1	-0.299949	0.22911409	-1.309	0.1970
X2	1	-1.389758	0.36134277	-3.846	0.0004
EDX1	1	0.002575	0.03330912	0.077	0.9387
EDX2	1	0.128380	0.05476328	2.344	0.0234

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	662.709146	.	.
EDUC	1	50.667770	0.39553836	0.93782707
X1	1	2.885000	0.03592084	0.33382461
X2	1	24.899674	0.24332686	0.35006598
EDX1	1	0.010058	0.00012988	0.32974433
EDX2	1	9.250621	0.10672017	0.33447852

Question 3 (12 points total)

When a child's birthday is in the late summer or early fall months, parents are often faced with the dilemma of whether to start the child in kindergarten as a very young (or not quite) 5-year-old or to wait a year. Conventional wisdom suggests that it is better to hold these children back, such that they are the oldest in their class, than to push them ahead and have them be the youngest in their class. Specifically, it is believed that being older is related to better social adjustment.

Education researchers are interested in this question, and also wonder if the effect of relative age on social adjustment persists as late as the fifth grade. They sample 10 first-grade classrooms and 10 fifth-grade classrooms. In each class, they identify three target children: the youngest child, the child with the median age, and the oldest child. All of the student in the 20 classrooms complete a questionnaire in which they are asked to rate the 3 target children in their class on a measure of social desirability. The dependent variable of interest is a composite social adjustment score (ADJUST) calculated for each target child. In other words, each of the three tarrget children in each of the classrooms has a score that is the mean social adjustment rating given that child by classmates. The primary questions of interest are: is there a linear effect of relative age, such that the youngest target child has poorer adjustment than the oldest? Does this difference depend on grade? On average, do first- and fifth-graders differ in mean level of social adjustment?

Set up a complete source table for the analyses suggested by the researchers' questions. Fill in the SOURCE and DEGREES OF FREEDOM columns. Make sure that your labels in the SOURCE row are descriptive enough for us to understand what that row represents (e.g., don't label a row "X1" unless you define X1).