

Name Key
(Please ALSO write your name on the last page of the quiz.)

Ag 400 - Quiz 4
Fall 2006

- A. Suppose you are interested in studying the relationship between the number of errors a student makes on a test and his/her evaluation of the "fairness" of that test.

Number of errors is (X) was measured by counting the number of wrong answers the student had on his/her test.

Evaluation of the fairness of the test was indexed by asking each student to rate on a scale of 0 to 10 the "fairness" of the test in assessing student knowledge.

The following data were compiled from a sample of 25 students:

	<u>Number of Errors (X)</u>	<u>Evaluation of test (Y)</u>
Mean Score	10	5
Sum of the Scores	250	125
Sum of the Squares of the Scores	2,900	725
Sum of the Squares of the Deviations about the Mean	400	100
Sum of the Cross Product Deviations ($\sum xy$)		-90

- (20) 1. Test the statistical significance of the relationship between number of errors made and evaluation of the test. Use the .05 level and a two-tailed test to determine statistical significance. Specify the following:

- a. State the Null Hypothesis and the Alternative Hypothesis.

H_0 : there is no linear relationship between Number of Errors and Evaluation of Test.

H_A : there is a linear relationship ...

- b. Show the necessary calculations.

$$r = \frac{-90}{\sqrt{(400)(100)}} = -.45$$

- c. Specify degrees of freedom.

$$df = 23$$

- d. Indicate the "critical value" of the test statistic.

$$r_{.05} = .396$$

- e. Reject or do not reject the Null Hypothesis.

Reject H_0

- f. What is the probability that you have made a Type I error in (e) above?

$$.01 < \alpha < .05$$

- g. What is the probability that you have made a Type II error in (e) above?

$$\beta = 0$$

- h. State your conclusion precisely in terms of the problem.

Number of errors students made on the test is negatively and linearly related to their evaluation of the fairness of the test.

- (10) 2. Estimate the Evaluation Score (Y) given by a student who made 10 errors on the test.

$$b_{yx} = \frac{-90}{400} = -.225$$

$$a_{yx} = \frac{125 - (-.225)(250)}{25} = 7.25$$

$$Y = 7.25 + (-.225)(10) = 5$$

- (10) 3. Estimate the Number of Errors made by a student who rated the test with an Evaluation Score of 8.0.

$$b_{xy} = \frac{-90}{100} = -.90$$

$$a_{xy} = 10 - (-.9)(5) = 14.5$$

$$X = 14.5 + (-.90)(8) = 7.3$$

- (10) 4. Indicate whether each of the following statements is true (T) or false (F) in terms of the above data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

T a. The statistical unit in this problem is a student.

T b. In this analysis b_{yx} and b_{xy} are both negative. When b_{yx} is negative, b_{xy} will always be negative in the same analysis.

T c. If a student made no errors on the test, using the above analysis, you would estimate that his/her rating of the fairness of the test (the Evaluation Score) would equal a_{yx} .

F d. Using the above analysis, for every 1 point increase in Evaluation Score, you would predict that the student's number of errors would decrease by b_{xy} percent.

F e. If you had access to the Number of Errors and the Evaluation Scores for each of the 25 students in the sample, the relationship between Number of Errors and Evaluation Score could have been tested for statistical significance using a paired t-test.

- (50) B. The attached computer output was obtained from data compiled from a random sample of 300 Mudville residents to assess the relationships of age and education to resident's knowledge of current local issues. The following variables were used: (Missing data are possible for all variables.)

Y = Knowledge Score (SCORE)

This variable measures each resident's knowledge of current events based on a series of 20 questions about local issues and happenings. The higher the person's score, the greater his/her knowledge.

X₁ = Age of respondent in years (AGE)

X₂ = Education of Respondents (ED)

Number of years of formal schooling that respondent completed

Indicate whether each of the following statements is true (T) or false (F) in terms of these data. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

- F 1. The dependent variable in this analysis is Knowledge Score and the statistical unit is a point on the Knowledge Score Scale.
- F 2. The figures on the output are parameters since they allow us to generalize to the population of Mudville residents.
- T 3. SCORE, AGE, and ED are all measured by interval scales for this analysis.
- F 4. In this data set, the mean for ED equals 13.76. This means that most people in the sample have completed more than high school (i.e. ^{made} have completed more than 12 years of formal schooling).
- T 5. The bivariate linear interrelationship between age and education is negative and statistically significant at the .001 level using a two-tailed test.
- F 6. Overall, for every one year increase in age you can predict a .184 point decrease in Knowledge Score.

- T 7. Based on the data given on Page 1, if you reject the following Null Hypothesis in favor of the following Alternative Hypothesis ($\alpha = .008, \beta = 0$)
- H_0 : In the population there is no linear relationship between education and Knowledge Score.
- H_A : There is a linear relationship between education and Knowledge Score in the population.
- T 8. If a subject in the sample did not indicate his/her age, he/she was omitted from all of the analysis presented here.
- F 9. The "R" reported at the top of Page 2 in the "Model Summary" table is the Multiple Correlation Coefficient. This R is statistically significant at the .001 level.
- F 10. Although SPSS has failed to print the minus sign for the R reported at the top of Page 2 of the output, this coefficient ($R = .205$) should be negative since the negative relationship of AGE to SCORE is stronger than the positive relationship of ED to SCORE.
- F 11. Approximately 20.5% of the total variation in Knowledge Scores is accounted for by the combined linear effects of ED and AGE.
- T 12. In the analysis, the "Adjusted R Square" reported in the Model Summary table at the top of Page 2 is smaller than R Square. Adjusted R Square is never greater than the corresponding R Square for the same analysis.
- T 13. If the combined linear effects of AGE and ED on SCORE actually equals zero in the population, the probability of obtaining a Multiple Correlation of .205 or more simply by chance in this sample was .002.
- F 14. Adjusting for the effects of ED, for every 1 year increase in AGE, Knowledge Scores would be estimated to decline by 8.055 points based on the analysis presented here.
- T 15. When the effect of AGE is controlled, the linear relationship of ED to SCORE is not statistically significant at the .05 level.

- T 16. In the bivariate analysis, the relationship of education to Knowledge Score was inflated. When the effect of AGE was controlled the relationship of ED to SCORE declined in strength.
- T 17. In the bivariate analysis, the relationship of age to Knowledge Score was inflated. When the effect of ED was controlled, the strength of the relationship of AGE to SCORE declined.
- T 18. The Zero-order Correlations on Page 2 for AGE and ED are both statistically significant at the .01 level.
- F 19. Adjusting for the effect of education, every one year increase in age is associated with .139 points decrease in Knowledge Score.
- T 20. Approximately 2% of the variation in Knowledge Scores in the sample that is not associated with variation in education is accounted for by variation in age of the subjects.
- F 21. If AGE were deleted from the regression model presented on Page 2 of the output, the correlation coefficient for the reduced model would not be statistically significant at the .05 level.
- T 22. If ED were deleted from the regression analysis on Page 2, the proportion of the variation in SCORES accounted for by the model declines from .042 to .034, but this decline is not statistically significant at the .05 level.
- T 23. Overall, for every one year increase in age, Knowledge Scores would be predicted to decline by .10 points.
- T 24. Based on this output, and using the most parsimonious predictive model presented, you should predict that a 44-year old Mudville resident would have a Knowledge Score of approximately 9.9.
- F 25. In the regression model presented on Page 3 of the output, the positive correlation of AGE to SCORE is statistically significant at the .001 level.

```

SAVE OUTFILE='C:\TEACHING\McrHw_05.sav'
/COMPRESSED.
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA ZPP
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT score
/METHOD=ENTER age ed .

```

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
SCORE	10.00	3.408	297
AGE	43.31	6.267	297
ED	13.76	5.712	297

Correlations

		SCORE	AGE	ED
Pearson Correlation	SCORE	1.000	-.184	.152
	AGE	-.184	1.000	-.376
	ED	.152	-.376	1.000
Sig. (1-tailed)	SCORE	.	.001	.004
	AGE	.001	.	.000
	ED	.004	.000	.
N	SCORE	297	297	297
	AGE	297	297	297
	ED	297	297	297

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	ED, AGE ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: SCORE

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.205 ^a	.042	.035	3.348

a. Predictors: (Constant), ED, AGE

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	143.914	2	71.957	6.421	.002 ^a
	Residual	3294.632	294	11.206		
	Total	3438.545	296			

a. Predictors: (Constant), ED, AGE

b. Dependent Variable: SCORE

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	12.700	1.718		7.392	.000			
	AGE	-8.055E-02	.034	-.148	-2.404	.017	-.184	-.139	-.137
	ED	5.730E-02	.037	.096	1.559	.120	.152	.091	.089

a. Dependent Variable: SCORE

Regression

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	AGE ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: SCORE

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.184 ^a	.034	.031	3.356

a. Predictors: (Constant), AGE

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	116.687	1	116.687	10.362	.001 ^a
	Residual	3321.859	295	11.261		
	Total	3438.545	296			

a. Predictors: (Constant), AGE

b. Dependent Variable: SCORE

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	14.339	1.362		10.529	.000			
	AGE	-.100	.031	-.184	-3.219	.001	-.184	-.184	-.184

a. Dependent Variable: SCORE