

Sociology 593
Exam 1
February 12, 1999

I. *True-False.* (20 points) Indicate whether the following statements are true or false. If false, briefly explain why.

1. As more independent variables are added to a model, adjusted R^2 will always increase or stay the same.
2. In a bivariate regression, Multiple R is always the same as the bivariate correlation between the IV and the DV.
3. The null and alternative hypothesis are

$$H_0: \beta = 1$$

$$H_A: \beta \neq 1$$

In the sample, $b = 1.6$, $s_b = .5$, $N = 10,000$. If we are using the .05 level of significance, we should reject the null.

4. $R^2 = .7$, $sr_3^2 = .2$, $sr_4^2 = .1$. Hence, if X3 and X4 are removed from the equation, R^2 will drop to .4.
5. Random measurement error can distort comparisons of variable effects.

II. *Short answer.* (15 pts. Each, 45 points total). For each of the following, indicate (i) what problem appears to be present (and how you can tell that from the information given) (ii) why you should be concerned about the problem, i.e. what harmful effects might it have when estimating regression models, and (iii) possible solutions. When discussing solutions, be sure to look carefully at the information presented; if, in this particular case, some solutions appear to be better than others, explain why.

1.

Descriptive Statistics

	Mean	Std. Deviation	N
Applicant income	\$47.11	\$35.77	988
Is Applicant black?	1.01E-02	.10	988
FEMALE	.1154	.3196	988

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12462.442	2	6231.221	4.910	.008 ^a
	Residual	1250105.311	985	1269.142		
	Total	1262567.753	987			

- a. Predictors: (Constant), FEMALE, Is Applicant black?
 b. Dependent Variable: Applicant income

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error				Beta	Lower Bound	Upper Bound	Tolerance
1	(Constant)	48.383	1.212		39.920	.000	46.005	50.761		
	Is Applicant black?	-21.183	11.331	-.059	-1.870	.062	-43.418	1.052	.999	1.001
	FEMALE	-9.164	3.550	-.082	-2.581	.010	-16.130	-2.198	.999	1.001

- a. Dependent Variable: Applicant income

Casewise Diagnostics^a

Case Number	Std. Residual	Applicant income	Predicted Value	Residual
1	26.941	\$999	\$39.22	\$959.78

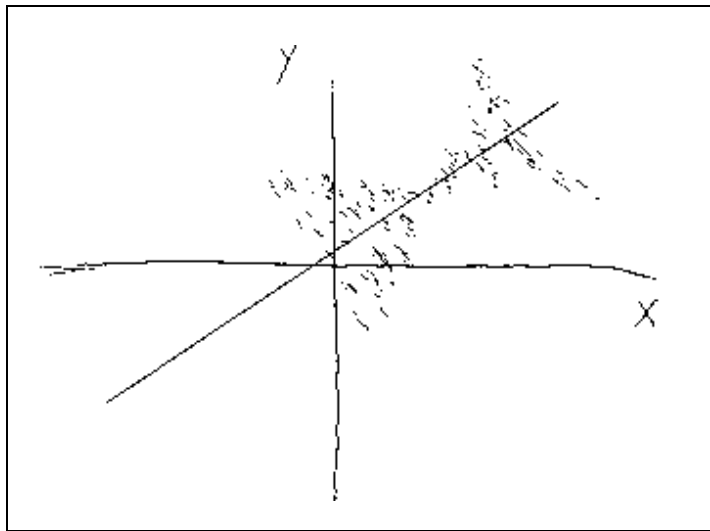
- a. Dependent Variable: Applicant income

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	\$27.20	\$48.38	\$47.11	\$3.55	988
Residual	-\$40.38	\$959.78	\$0.00	\$35.59	988
Std. Predicted Value	-5.603	.358	.000	1.000	988
Std. Residual	-1.134	26.941	.000	.999	988

- a. Dependent Variable: Applicant income

2. A researcher is interested in the relationship between income, gender, and racial prejudice. A sample of 1,000 men and 1,000 women is interviewed. All respondents report their gender and complete a 10 item scale that measures racial prejudice. However, 50 men and 200 women do not answer the question on income.
3. A researcher is interested in the effect of socio-economic status (X) on political ideology (Y). When she plots X against Y, she gets the following:



III. Computation and interpretation. (35 points; up to 10 points extra credit)

In their classic 1985 paper, "Ability grouping and contextual determinants of educational expectations in Israel," Shavit and Williams examined the effect of ethnicity and other variables on the achievement of Israeli school children. There are two main ethnic groups in Israel: the Ashkenazim - of European birth or extraction - and the Sephardim, most of whose families immigrated to Israel during the early fifties from North Africa, Iraq, and other Mid-eastern countries.

The variables used here are:

X1 - Ethnicity (SPHRD) - a dummy variable coded 1 if the respondent or both his parents were born in an Asian or North African country, 0 otherwise

X2 - Parental Education (PARED) - A scale which ranges from a low of 0 to a high of 1.697

X3 - Scholastic Aptitude (APTD) - A composite score based on seven achievement tests.

Y - Grades (GRADES) - Respondent's grade-point average during the first trimester of eight grade. This scale ranges from a low of 4 to a high of 10.

We will look at two models that can be estimated using these data. In both models, GRADES is the dependent variable.

MODEL I: Bivariate Regression

Descriptive Statistics

	Mean	Std. Deviation	N
Sephardim	.44000	.50000	10609
Parental Education Scale	.82000	.46000	10609
Scholastic Aptitude	6.46000	2.11000	10609
Grade Point Average	7.12000	1.42000	10609

Correlations

		Sephardim	Parental Education Scale	Scholastic Aptitude	Grade Point Average
Pearson Correlation	Sephardim	1.000	-.590	-.460	-.260
	Parental Education Scale	-.590	1.000	.530	.330
	Scholastic Aptitude	-.460	.530	1.000	.720
	Grade Point Average	-.260	.330	.720	1.000

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.260 ^a	.068	.068	1.37123

a. Predictors: (Constant), Sephardim

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1445.962	1	1445.962	769.019	.000 ^a
	Residual	19944.009	10607	1.880		
	Total	21389.971	10608			

a. Predictors: (Constant), Sephardim

b. Dependent Variable: Grade Point Average

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	7.445	.018		419.808	.000
	Sephardim	-.738	.027	-.260	-27.731	.000

a. Dependent Variable: Grade Point Average

a. (10 pts.) Interpret the results from Model I. What percentage of the students are Sephardim? Which group has the higher grade point average? Based on this, which group do you think is considered "disadvantaged" in Israel?

Model II: Multivariate Regression

Model Summary

R Square	[1]
----------	-----

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11241.697	3	3747.232	[2]	.000 ^a
	Residual	10148.275	10605	.957		
	Total	21389.971	10608			

a. Predictors: (Constant), Scholastic Aptitude, Sephardim, Parental Education Scale

b. Dependent Variable: Grade Point Average

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	3.756	.042		89.758	.000		
	Sephardim	.212	.024	.075	8.790	.000	.622	[3]
	Parental Education Scale	-.111	.027	[4]	-4.037	.000	.567	1.763
	Scholastic Aptitude	[5]	.005	.773	95.746	.000	.686	1.458

a. Dependent Variable: Grade Point Average

b. (20 pts.) Fill in the missing information (1) - (5).

c. (5 pts.) According to Model II, if you are testing

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 < 0$$

and you are using the .01 level of significance, should you accept or reject the null hypothesis?

d. (Optional; 10 pts. Extra credit) In model I, the effect of Sephardim is negative. In model II, the effect is positive. Offer a substantive explanation (or theory) as to why this happens.