Soc 3811 Basic Social Statistics
Third Midterm Exam Spring 2010

Your Name [50 points]: ______________________________ ID #:______________

Your TA: Kyungmin Baek_____ Meghan Zacher_____ Frank Zhang_____

INSTRUCTIONS:

(A) Write your name on the line at top front of every sheet.

(B) If you use a page of notes in taking this exam, sign & insert it inside this booklet before turning in your exam.

(C) *Show your calculations for numerical problems in the space provided!*

- -

1. Using a random sample of 436 Vietnamese rice farmers, an agronomist regresses the number of bushels harvested per acre (Y) on pounds of fertilizer per acre (X), resulting in this prediction equation (the standard errors are in parentheses):

$$\hat{Y}_i = \underset{(17.4)}{53.6} + \underset{(1.2)}{3.3} X_i \qquad R^2_{YX} = 0.457$$

Use her equation to estimate how many bushels of rice were harvested by farmers using these amounts of fertilizer **[5 points]:**

X_i	$\hat{Y}_i$
10	________
15	________
20	________
25	________
30	________

2. Calculate the coefficients of determination (R^2) for three different regression equations with these sums of squares and numbers of independent variables **[5 points]:**

(1): Four independent variables

SS_{ERROR} = 18,271

SS_{TOTAL} = 24,482

R^2 = ______________

(2): Eight independent variables

$SS_{REGRESSION}$ = 759

SS_{ERROR} = 3,824

R^2 = _______________

(3): Three independent variables

$SS_{REGRESSION}$ = 4,263

SS_{TOTAL} = 8,327

R^2 = ______________

Your name:____

3. Test a family sociologist's research hypothesis that more years of education a mother has (X), the fewer the fights among her children (Y). This bivariate regression equation used a sample of 438 mothers with two children between ages 6 to 12 years (the standard errors are in parentheses):

$$\hat{Y}_i = \underset{(8.2)}{48.7} \quad \underset{(0.6)}{-1.4}\, X_i \qquad\qquad R^2_{YX} = 0.046$$

Write his null and research hypotheses about β_{YX} in symbolic form; show your calculations of the test statistic; state your decision about H_0 and the lowest probability of Type I (false rejection) error. State the substantive conclusion of your decision. (See the critical values tables on page 11 of this exam.) **[5 points]:**

H_0: ______________________________

H_1: ______________________________

Decision about H_0: ____

Probability of Type I error: ____________________________________

State the substantive conclusion of your decision:

____

____

4. Now test the null hypothesis about the coefficient of determination from the bivariate regression in the preceding problem. Write the null and research hypotheses in symbolic form; show your calculations of the test statistic in the ANOVA table; state your decision about H_0 and the lowest probability of Type I error. State the substantive conclusion of your decision. (Critical-values table is on page 11). **[5 points]:**

H_0: ______________________________

H_1: ______________________________

Source	SS	df	MS	F
Regression	**791**			
Error	**16,409**			
Total	**17,200**		**- - - - - - - - - - - - - - - -**	

Decision about H_0: ____

Probability of Type I error: ____________________________________

State the substantive conclusion of your decision:

____

Your name:____

5. A political scientist analyzes congressional elections by regressing the vote for Republican candidates on eight independent variables in a sample of 215 electoral districts. Her estimated $R^2 = 0.374$. Show your computation of $R^2_{adj.}$. **[5 points]:**

$R^2_{adj.} =$ ____

6. An economist hypothesizes that unemployment in the Great Recession can be explained by six corporate variables. She regresses a measure of job layoffs on those six independent variables for a sample of 215 corporations, producing the ANOVA table below. Write her null and research hypotheses about the coefficient of determination in symbolic form; complete the analysis of variance table; state your decision about H_0; give the <u>lowest probability</u> of Type I error; and state your substantive conclusion. **[5 points]:**

H_0: ________________________________

H_1: ________________________________

Source	**SS**	**df**	**MS**	**F**
Regression	**556**			
Error	**9,457**			
Total	**10,013**		- - - - - - - - - - - - - - -	

Decision about H_0: ____

Probability of Type I error: ____

State the substantive conclusion of your decision:

____

____

Your name:____

7. A real estate researcher hypothesizes that both house and neighborhood factors affect home prices. For a sample of 547 homes, he first regresses home price on six house variables and finds a coefficient of determination $R^2 = 0.253$. After adding four neighborhood independent variables to the equation, the second equation $R^2 = 0.297$. Write the researcher's null and alternative hypotheses in symbolic notation; carry out the appropriate statistical test; state your decision about H_0; and, if you reject H_0, report the lowest probability of Type I error. **[5 points]:**

H_0: ____

H_1: ____

Decision about H_0: ____

Probability of Type I error: ______________________________________

State the substantive conclusion of your decision:

____

____

8. Using data from a sample of 673 released felons, a criminologist regresses a measure of recidivism (Y) on education (X_1), hours worked (X_2), and age (X_3), producing this unstandardized prediction equation:

$$\hat{Y}_i = 16.48 - 0.41\,X_{1i} - 0.28\,X_{2i} - 0.19\,X_{3i} \qquad R^2_{adj.} = 0.327$$
$$(5.36) \quad (0.09) \qquad (0.14) \qquad (0.04)$$

Use the standard deviations below to change the unstandardized b's in the equation above into standardized coefficients (β*). Then write the standardized regression equation and identify the strongest predictor of recidivism. **[5 points]:**

VARIABLE	**STD. DEV.**
Recidivism (Y)	**6.0**
Education (X_1)	**4.0**
Employment (X_2)	**8.0**
Age (X_3)	**15.0**

Standardized Eq.: ______________________________

Strongest predictor: ______________________________

Your name:____

9. A Minnesota historian uses 28,472 Census records from 1890 to study how education (Y_i) varied among European immigrant groups. The nonordered variable ETHNIC has five categories (below). Show the coding scheme to change ETHNIC into a set of dummy variables for use as independent variables in a multiple regression equation. **[5 points]:**

ETHNIC					
1. Swedish					
2. Norwegian					
3. German					
4. Irish					
5. Other					

Use this regression equation to estimate the education of immigrants with each ethnicity:

$$\hat{Y}_i = 9.4 + 1.2\, D_{Swedish} - 0.8\, D_{Norwegian} + 1.5\, D_{German} - 1.2\, D_{Irish} \qquad R^2_{adj.} = 0.246$$

$\hat{Y}_{\text{Swedish}}$ = ____________

$\hat{Y}_{\text{Norwegian}}$ = ____________

$\hat{Y}_{\text{German}}$ = ____________

$\hat{Y}_{\text{Irish}}$ = ____________

$\hat{Y}_{\text{Other}}$ = ____________

10. A religion scholar regresses the frequency of prayer on: a 10-point religiosity measure (X_{RELIG}); age (X_{AGE}), coded in years; and a race dummy variable (D_{RACE}), coded 1 = nonwhite, 0 = white. The unstandardized and standardized ANCOVA equations are:

$$\hat{Y}_i = 35.4 \quad + 2.3\, X_{RELIG} \quad + 0.3\, X_{AGE} + 4.7\, D_{RACE} \qquad\qquad R^2_{adj.} = 0.418$$

$$\hat{Z}_{Y_i} = \quad + 0.5\, Z_{RELIG} \quad + 0.3\, Z_{AGE} + 0.2\, Z_{RACE}$$

Calculate the predicted frequency of prayer for:

(a) A 20 year old nonwhite person of low religiosity ($X_{RELIG} = 2$):

$\hat{Y} =$ ____________________

(b) A 70 year old white person of high religiosity ($X_{RELIG} = 8$):

$\hat{Y} =$ ____________________

Write a brief substantive interpretation of how the three predictors each affect frequency of prayer, and indicate which predictor(s) has the largest effect:

____

____

____

Critical values (c.v.) of Z and *t* for large samples

α	One-tail c.v.	Two-tail c.v.
.05	1.65	$\pm$1.96
.01	2.33	$\pm$2.58
.001	3.10	$\pm$3.30

Critical values (c.v.) of F distributions for large samples

df_R, df_E	α = .05	α = .01	α = .001
1, ∞	3.84	6.63	10.83
2, ∞	3.00	4.61	6.91
3, ∞	2.60	3.78	5.42
4, ∞	2.37	3.32	4.62
5, ∞	2.21	3.02	4.10
6, ∞	2.10	2.80	3.74
7, ∞	2.01	2.64	3.47
8, ∞	1.94	2.51	3.27
9, ∞	1.88	2.41	3.10
10, ∞	1.83	2.32	2.96