S370 HOMEWORK 6 **Spring 2010** **PROF. W. E. BECKER**

Regression Analysis: Ex 12.20, 13.9, 13.13, 13.23a, 13.36, 13.37, 13.38, 14.2, 14.8, & 14.11. (10 pts each)

12.20

a. The slope coefficient is 0.5019 and the intercept coefficient is 20.8247.As a result, we have the regression equation $\hat{y}$ = 20.8247 + 0.5019 x.

b. = 20.8247 + 0.5019(40) = 40.9007

13.9

a. The coefficient of determination is r^2 = 0.91429. This is calculated to be the ratio of RegSS over TSS.

b. The standard error of the coefficient of X, b, is 0.06999 and is calculated via the formula seen on p474.

c. The significance level = 0.00012 on the regression output is the (two-tail) p-value used to test the hypothesis H_0: Intercept = 0 vs. H_A: Intercept ≠ 0; it is determined in EXCEL by =TDIST(14.832,4,2). This small p-value implies that the intercept is significant in the population.

d. The (one-tail) p-value 0.00142 for a test H_0: $\beta = 0$ vs. H_A: $\beta < 0$ implies that the alternative hypothesis H_A: $\beta < 0$ is accepted at the 0.01 Type I error level. We can conclude that there is a significant negative relationship between X and Y.

13.13

a. EXCEL gives the following output:

Regression Statistics	
Multiple R	0.744301656
R Square	0.553984955
Adjusted R Square	0.529206342
Standard Error	0.724780386
Observations	20

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	0.008237232	0.233033818	0.035348	0.972191
Clinical Trials	0.219934102	0.046513829	4.728359	0.000168

The least-squares regression equation is NUMBER-IN-MARKET = 0.008+ 0.2199(NUMBER-IN-TRIALS). That is, from the slope coefficient we know that for an increase of 10 clinical trial, the predicted number of drugs for human use on the market increases by 2.199. For no clinical trials the predicted number of drugs for human use is 0.008.

b. This relationship is statistically significant down to a Type I error level of 0.000168.

13.23

a. The slope coefficient is significant at the 0.05 Type I error level. (In fact, the null hypothesis in a two-tail test could be rejected down to the 0.02928 Type I error level.) Thus, there is a relationship between days of school and math scores.

13.36

a. The slope coefficient 0.5317 implies that a one percentage point increase in industrial production is associated with a 0.5317 percentage point increase in employment.

b. We need to test the hypothesis H_0: $\beta = 0$ vs. H_A: $\beta \neq 0$ where β is the population slope parameter. At a typical α level 0.05, we can reject H_0 because the calculated t 4.895 is greater than the critical $t_{(0.025,df=5)}$ = 2.571, or because the p-value 0.00449 is less than $\alpha = 0.05$. We can conclude that there is a significant relationship between a change in industrial production and change in employment in the population.

13.37

When the percentage change in production is 3.10, the percentage change in employment is predicted to be 0.9 [= -0.7046 + 0.5317(3.10)].

13.38

TotalSS = RegressionSS + ErrorSS = 4.6286 + 0.9657 = 5.5943

(ResidualSS) ErrorSS = 0.9657

(The coefficient of determination) R^2 = .8274.

This R^2 value indicates that the 82.74 percentage of the variability of employment around its sample mean (as reflected in the total sum of squares) is explained by the regression.

14.2

a. If a jockey's experience increases by one year, then the predicted horse speed increases or decreases by b miles per hour (depending on the sign of b), holding all other things (jockey's sex, weight carried on the horse and horse's condition) fixed.

b. When a jockey is male (sex = 1), then the predicted horse speed is c miles per hour faster or slower (depending on the sign of c) than with a female jockey, holding all other things (jockey's experience, weight carried on the horse and horse's condition) fixed.

14.8

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.37778
R Square	0.142718

Adjusted R Square	-0.14304
Standard Error	266.9198
Observa	13

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	106747.6182	35582.54	0.499430964	0.691873741
Residual	9	641215.4587	71246.16		
Total	12	747963.0769			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	2719.540554	505.0974326	5.38419	0.000442
ROOMS	-4.940269143	53.24033838	-0.09279	0.928102
TOTAL	-0.125847947	0.125863244	-0.99988	0.343492
AGE	-1.337208767	3.584004988	-0.3731	0.717707

a. The least squares regression equation is
RICE = 2719.54 - 4.940272 ROOMS - 0.125848 TOTAL - 1.33721 AGE,
where TOTAL = MAIN FL + UPPER FL + BASEMENT

b. At ROOMS = 8, TOTAL = 2500 and AGE = 0, the predicted price is:

RICE = 2719.54 - 4.940272 (8) - 0.125848 (2500) - 1.33721 (0) = $236,540

14.11

a. At SALES = 40 and PRODUCT = 20, the predicted vale of y (net income) is
INCOME = -17.793 + 0.9121(40) + 0.1589(20) = $218,690

b. The coefficient 0.9121 indicates that an additional salesperson is associated with a $9,121 increase in predicted net income, holding the number of PRODUCTS constant.