STA 110B Name_____

Spring 2000 Section_____

Quiz 11

week of 17APR2000

Below you will find the SAS INSIGHT output for a multiple regression performed on a data set of 49 patients, all of whom had the same diagnosis. The response variable is the natural logarithm of the hospital charges incurred by the patients; the regressors are the age of the patient, the age of the patient squared, and the gender of the patient. (The variable IS_MALE equals 1 if the patient is male and 0 if female.)

Summary of Fit

Mean of Response	9.4745	R-Square	0.6431
Root MSE	0.4536	Adj R-Sq	0.6193

Parameter Estimates

Variable	DF	Estimate	Std Error	T Stat	Prob >|T|	Tolerance	Var Inflation
INTERCEPT	1	6.2068	0.5817	10.6694	0.0001	.	0
AGE	1	0.0729	0.0186	3.9211	0.0003	0.0392	25.4908
IS_MALE	1	-0.0033	0.1366	-0.0243	0.9807	0.9002	1.1108
AGE_SQ	1	-0.0003	0.0001	-2.2471	0.0296	0.0395	25.3394

Note: There was a lot of confusion due to unclear printing of the regression output. Many students mistook "3"s for "9"s. For this reason, I will have written the answers for both readings ("9"s or "3"s) of the output.

1. (2 points) Write out the fitted model for the logarithm (hereafter written just "log") of hospital charges incurred as a function of patient's age, gender, and age squared.

$$\hat{Y} = 6.2068 + 0.0729 * AGE - 0.0033 * ISMALE - 0.0003 * AGE_SQ$$

or

$$\hat{Y} = 6.2068 + 0.0729 * AGE - 0.0099 * ISMALE - 0.0009 * AGE_SQ$$

where $\hat{Y}$ equals predicted log charges.

2. (2 points) What amount of charges would you expect using this model for a patient who is 50 years old and female?

Log charges:

$$\hat{Y} = 6.2068 + 0.0729 * 50 - 0.0033 * 0 - 0.0003 * 50 * 50 = 9.1018$$

or

$$\hat{Y} = 6.2068 + 0.0729 * 50 - 0.0099 * 0 - 0.0009 * 50 * 50 = 7.6018$$

To get back to charges (not on log scale), we need to take the exponent, yielding the answer(s):

$$e^{9.1018} \approx 8971$$

or

$$e^{7.6018} \approx 2002$$

3. (2 points) Estimate the expected change in log charges associated with the patient being male. (Make sure to specify whether this is an increase or a decrease.)

0.0033 decrease in log charges or 0.0099 decrease in log charges

4. (2 points) Does a patient's gender appear to have a significant effect on the log charges incurred? State your reasoning briefly in mathematics or a sentence.

The p-value associated with the gender of the patient is 0.9807/0.3807, so it doesn't seem that gender has a significant effect on the charges incurred. Remember this p-value the result of the test of these two hypotheses: $H_0 : \beta_{ISMALE} = 0$ $H_A : \beta_{ISMALE} \neq 0$
So, this p-value doesn't allow us to reject the null hypothesis.

5. (2 points) What percentage of the total sum of squares of deviation (total SS) is explained by all the regressors?

$R^2 = 0.6431 = 64.31\%$ or $R^2 = 0.6491 = 64.91\%$