S370 Spring 2010 Final Exam key

I. (70 pts) An article in the *Wall Street Journal* "Artful Deal Advisers Look Beyond Quantity"stated:

"Take Merrill Lynch & Co. It was ranked No. 2 in the global merger league table (below), advising on $218 billion of completed deals so far this year . . . But when it came to fees for its work, Merrill, the second biggest brokerage firm, wound up in sixth place, earning $536.1 million in advisory fees . . . By comparison, rival Credit Suisse First Boston, which was not as highly ranked in terms of completed advisory volume, took second place with $638.2 million in fees so far this year. … 'It is not just the deal assignment per se, but the quality of your role that is driving your fees,' says Don Meltzer."

ADVISOR NAME	Fees (in millions)	Deal Value (in billions)
Goldman Sachs	847.32	322.50
Credit Suisse First Boston	638.21	130.37
Morgan Stanley	591.99	173.47
Citigroup	587.64	184.63
JP Morgan	545.97	182.60
Merrill Lynch	536.09	218.00
UBS	471.70	127.21
Deutsche Bank	392.89	111.04
Lehman Brothers	390.96	121.69
Lazar	347.25	175.50

A&B. What is the sample mean and variance of fees (make sure to specify unit of measurement)?
Mean is 535.002 millions of dol. and sample variance is 21615.13 millions of squared dol.

C&D. What is the estimated standard error of the mean of fees and standard deviation of the mean? **46.492075 mil. Dol**.

E&F. What are pint estimate and a 95 percent confidence estimate of the expected fees?
535.002 ±105.17238 mil. Dol.

G. Write the model, including all assumptions, to test for the existence of a linear relationship in which deal value might affect fees. (Make sure to define notation.)

Let y_i be the observed fee, x_i be the deal value, and u_i be the random error for ith firm, where $u_i \sim N(0,\sigma_u)$. Alpha is the y intercept and Beta is the slope for the relationship between $E(y_i \mid x_i)$ and x_i. That is, the model is $y_i = \alpha + \beta x_i + u_i$

H. H. Using the data provided and your assumed model in part G, what is the sample least-squares relationship between the deal value and fees? (Make sure to define notational symbols in the equation you write.) State explicitly what each number tells you.

$\hat{y}_i = 226.678 + 1.765x_i$ **, where** $\hat{y}_i$ **is predicted fees for ith firm.**

I

If *x (deal value)* = 0, then the predicted fees are 226,678 million dollars.
If deal value increased by one billion dollars, then predicted fees increased by 1.765 million dollars.

I. What is the coefficient of determination for your regression and what does it indicate? **R^2=0.559846031; that is, 55.98% of variability in y around its sample mean is explained by this regression**.

J&K. If the deal value is 200 billion dollars, what is the predicted dollar value of fees? Why does this predicted value differ from the point estimator in part E? **579.6513198 mil. Dol. The estimate of the mean in part E is an unconditional mean; it is not specific to any specific value of x. The expected value estimate in part J is condition on a specific x value.**

L. State the null and alternative hypotheses to test if there is no relationship between deal value and fees collected versus the idea that advising on deals of higher value increases fees collected. **H_o: Beta=0 versus H_a:Beta>0.**

M. What is the estimated standard error for the slope estimator? **0.55326667**

N. What is the p-value for your test in part I using your regression? **0.0064**

O. If the probability of a Type I error is set at 0.01, what conclusion can you draw and why? **Reject H_o because the p-value is smaller than the Type I error probability level.**

II. (30 pts) For the next 10 questions mark the most correct answer.

1. In general, which of the following is true about the least squares residuals? (e_i is the ith residual for a simple linear regression)

   A. $e_i = 0$ for all i. B. $\Sigma e_i \neq 0$ ***C***. $\Sigma e_i = 0$
   D. $e_1 = e_2 = e_3 = e_4 =. . . = e_n$; that is, all residuals are equal.

2. If the sample correlation coefficient is 0.8, this means that:

   A. 80 percent of the change in the dependent variable is explained by the independent variable (i.e, the regression).
   B. 64 percent of the variation in the dependent variable is explained by the regression.
   C. 64 percent of the change in the dependent variable is explained by the regression.
   D. 80 percent of the variation in the dependent variable is explained by the regression

3. The least squares regression equation is estimated by EXCEL so as to

A. minimize the sum of distances between the observed and predicted values of the dependent variable.
B. intersect as many of the observed values as possible.
C. minimize the distance between the mean and the actual values of the dependent variable.
D. minimize the sum of squared distances between the observed and predicted values of the dependent variable.

4. A human resource specialist estimated the relationship between salaries paid to executives in her firm and the years they were employed and their sex. The relationship estimated was:

$\hat{Y} = 20{,}000 + 999X_1 + 876X_2$, where Y = dollar salary, X_1 = years employed and X_2 = 1 if male and 0 if female.

What is the correct interpretation of these results?

A. Females and males with like seniority are predicted to receive $21,875.
B. Being a female, versus a male with like seniority, is "worth" $876.
C. Being a male, versus a female with like seniority, is "worth" $876.
D. Females and Males with no seniority can be expected to receive $20,000.

5. A larger simple random sample is preferred to a smaller one because the

A. larger sample will appear more normal.
B. larger sample will look more like the population.
C. sampling distribution of an estimator calculated from the larger sample will have a smaller variance.
D. The expected value of the estimator calculated from the larger sample will have a smaller variance and be closer to the parameter being estimated.

6. If two random variables X and Y have a negative covariance, then:

A. high values of x tend to be associated with high values of y and low values of x tend to go with low values of y.
B. high values of x tend to be associated with low values of y and low values of x tend to go with high values of y.
C. negative values of x tend to be associated with negative values of y, and vice versa.
D. the expected value of x times y is less than zero.

7. The prices of CD players are normally distributed, with a standard deviation of $12. A random sample of 16 CD player prices is drawn. A confidence interval for the mean price shows a ∀$9.30 margin of error. What was the level of confidence?

A. 99.9% **B.** 99.8% C. 99.7% D. 99.6%

8. For H_o: β(population slope) = 15 and H_a: β ≠ 15,, testing at the 10% level of significance, the probability of rejecting the null hypothesis when it is true is:

A. 0.1 and the probability of committing a Type I error is 0.1.
B. 5% and the probability of committing a Type II error is 95%.
C. 0.1 and the probability of committing a Type I error is 0.05.
D. 0.05 and the probability of committing a Type II error is 0.1.

9. What is the confidence level associated with a point estimator?

A. 100% B. 50% C. 0.333% **D.** 0%

10. If one does not reject the null hypothesis H_o: μ=20, and thus does not accept the alternative H_a: μ < 20, with $\bar{x}$=10, then

A. at the same significance level, the null hypothesis H_o: μ=0 can not be rejected in favor of the alternative H_a: μ>0, for $\bar{x}$=10.
B. a mistake has been made because the null hypothesis should be H_o: $\mu \geq 20$, if the alternative hypothesis is H_a: μ < 20.
C. there is statistical proof that μ is 20 and that it is not less than or greater than 20.
D. there is statistical proof that μ is 20 and that it is not less than 20.