**Math 644, Fall 2012**
**Homework 2 Due: Friday, 10/5/2012**

1. (Grade Point average): Refer to the Grade Point average problem in HW1.
   (a) Obtain 99% confidence interval for $\beta_1$. Does it include 0? Why are we interested in whether the interval include 0?
   (b) Test, using t-statistic, whether there is linear association between ACT and GPA when $\alpha = 0.01$ is used?
   (c) What is the $p$-value for part (b)? How does it support the conclusion in part (b)?
   (d) Obtain a 95% percent interval estimate of the mean freshman GPA for students whose ACT test score is 28. Interpret your confidence interval.
   (e) Mary Jones obtained a score of 28 on the entrance test. Predict her freshman GPA using a 95% prediction interval. Interpret your prediction interval.
   (f) Is the prediction interval in part (b) wider than the confidence interval in part (a)?

2. (Grade point average): Refer to the Grade Point average problem in HW1.
   (a) Set up the ANOVA table.
   (b) Conduct an F-test for $H_0 : \beta_1 = 0$ with $\alpha = 0.01$ is used?
   (c) What is the absolute magnitude of the reduction in the variation of $Y$ when $X$ is introduced into the model? What is the relative reduction? What is the name for the later measure?
   (d) Obtain $r_{XY}$ and attach the appropriate sign.
   (e) Which measure $R^2$ or $r$ has more clear-cut operational interpretation, explain.

3. Suppose we have the fitted regression function, $\hat{Y} = 350.7 - 1.4X$, where the $p$-value for the slope is 0.91. A student concludes that "the message I get here is that the more X is, the fewer Y will be". Comment.

4. An analyst fitted normal error regression model and conducted an F-test of $\beta_1 = 0$ versus $\beta_1 \neq 0$. The $p$-value of the test was 0.033 and the analyst concluded $H_a$ :

$\beta_1 \neq 0$. Was the $\alpha$ level used by the analyst greater than or smaller than 0.033? If the $\alpha$ level had been 0.01, what would be the appropriate conclusion?

5. (Airfreight breakage): A substance used in biological and medical research is shipped by airfreight to users in cartons of 1000 ampules. In the data, $X$ is the number of times the carton was transferred from one aircraft to another over the shipment route, and $Y$ the number of ampules found to be broken upon arrival. A linear regression model $Y = \beta_0 + \beta_1 X + \varepsilon$ is used.

   (a) Write down the estimated regression model. Plot the estimated regression function and the data. Does the linear regression function appear to give a good fit here?

   (b) Obtain a point estimator of the expected number of broken ampules when $X = 1$.

   (c) Estimate $\beta_1$ with 95% confidence interval. Interpret your interval estimate.

   (d) Conduct a t-test to decide whether or not there is a linear association between number of times a carton is transferred $X$ and the number of broker ampules $Y$ with significant level $\alpha = 0.05$? What is the $p$-value of the test?

   (e) The $\beta_0$ here is the mean number of ampuls broken when no transfer of shipment are made. Obtain 95% con?dence interval for $\beta_0$ and interpret it.

6. (Airfreight breakage): Refer to the Airfreight breakage problem.

   (a) Set up the ANOVA table. Which elements are additive?

   (b) Conduct F-test to check the linear association between $X$ and $Y$ with $\alpha = 0.05$.

   (c) Conduct t-test to check the linear association between $X$ and $Y$ with $\alpha = 0.05$. Compare with the F-test and show the equivalence.

   (d) Calculate $R^2$ and $r_{XY}$. What proportion of the variation in $Y$ is accounted for by introducing $X$ into the model?