Statistics 512: Homework#2
Solutions

1. For conducting statistical tests concerning the parameter β_1, why is the t test more versatile than the F test?

   Solution: The t-test is more versatile, since it can be used to test a one-sided alternative.

2. When testing whether or not $\beta_1 = 0$, why is the F test a one-sided test even though H_a includes both $\beta_1 < 0$ and $\beta_1 > 0$? [*Hint*: Refer to (2.57).]

   Solution: The F-ratio is large when $\beta_1^2 > 0$, which is the same as testing whether $\beta_1 > 0$ or $\beta_1 < 0$.

The next 5 problems continue the analysis of the plastic hardness data begun in the first homework.

3. Plot the data using `proc gplot`. Use `frame` as an option with the `plot` statement and include a smoothed function on the plot by using the `i = smnn` option on the `symbol1` statement, where nn is a number between 1 and 99. Is the relationship approximately linear?

   Solution: Yes, the relationship is reasonably linear. (See Figure 1.) There is some slight curvature at lower values of time but nothing substantial.

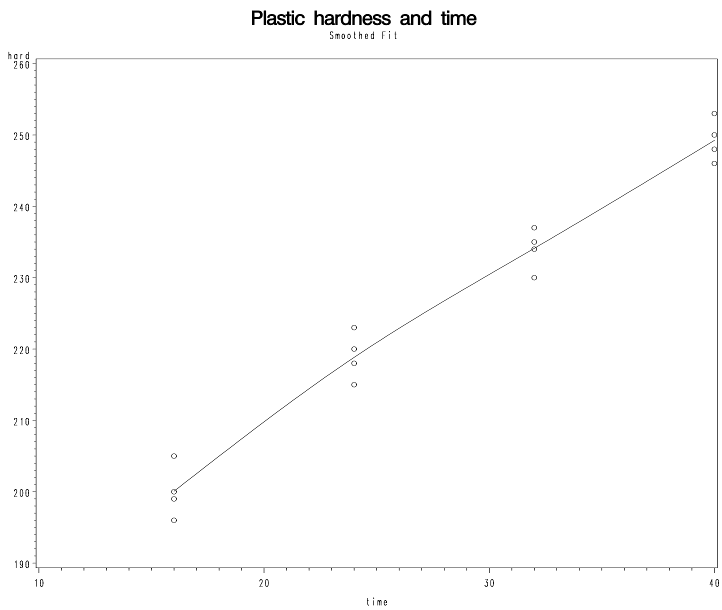

Figure 1: Graph for Problem 3

4. Plot the 95% bounds (confidence band) for the mean (use `i=rlclm` on the `symbol1` statement).

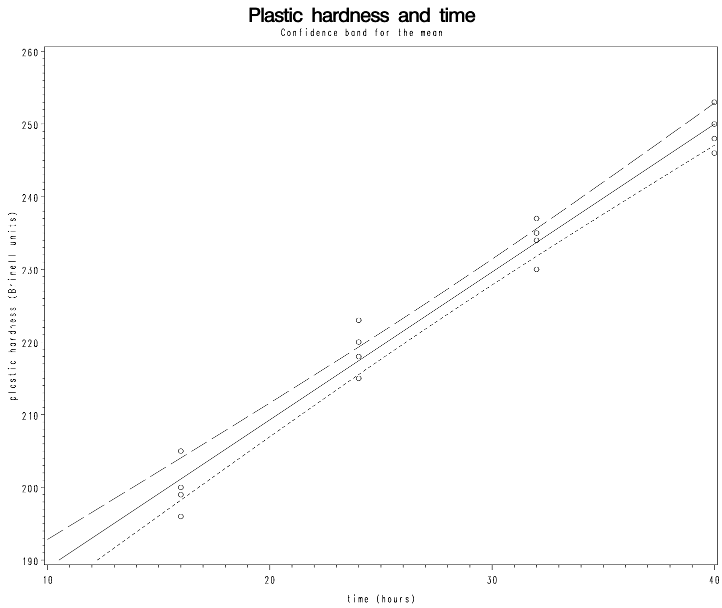

Figure 2: Graph for Problem 4

Solution: See the attached graph (Figure 2).

5. Plot the 95% bounds for individual observations (using `i=rlcli`).

   Solution: See the attached graph (Figure 3).

6. Give an estimate of the *mean* hardness that you would expect after 36 and 43 hours; and a 95% confidence interval for each estimate. Which confidence interval is wider and why is it wider?

   Solution: Based on the SAS output, the predicted value of hardness at 36 hours is 241.8, and a 95% confidence interval for the mean hardness is $[239.5, 244.2]$. The SAS output also gives that at 43 hours, the predicted hardness is 256.1, and a 95% confidence interval for mean hardness is $[252.7, 259.5]$. The confidence interval for $X = 43$ is wider because the value 43 is farther away from the sample mean $\bar{X}$ than is 36. As a result, the standard error for the prediction is larger.

7. Give a prediction for the hardness that you would expect for an *individual* piece of plastic after 43 hours; give a 95% prediction interval for this quantity.

   Solution: We predict a hardness of 256.1 after 43 hours; we are 95% confident that the hardness value will fall in the interval $[248.4, 263.4]$.

8. Calculate power for the slope using the results of text Problem 1.22 as follows. Assume $n = 16$, $\sigma^2 = MSE$, and $SS_X = 1280$.

   (a) Find the power for rejecting the null hypothesis that the regression slope is zero using an $\alpha = 0.05$ significance test when the alternative is $\beta_1 = 0.5$.

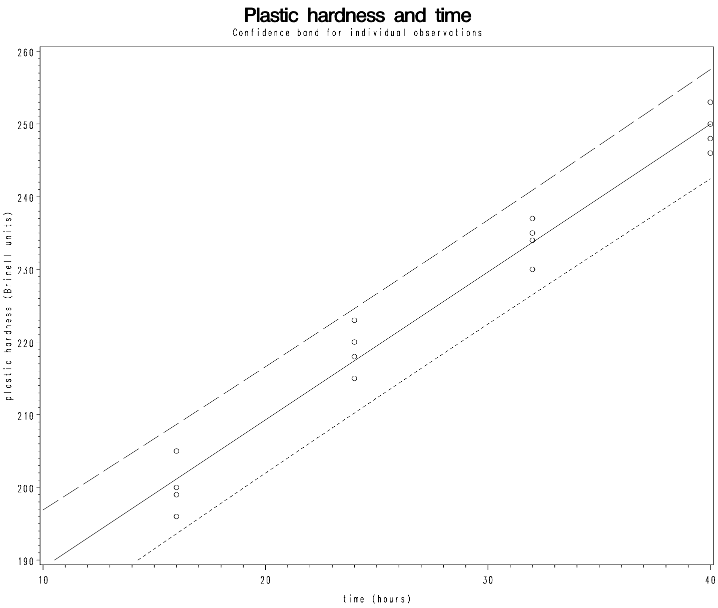

Figure 3: Graph for Problem 5

Solution: The power against $H_A : \beta_1 = 0.5$ (calculated using SAS) is 0.99922 .

(b) Plot the power as a function of β_1 for values of β_1 between -2.5 and +2.5 in increments of 0.25 .

Solution: See the attached graph (Figure 4).

9. Given that $R^2 = SSM/SST$, it can be shown that $R^2/(1 - R^2) = SSM/SSE$. If you have $n = 28$ cases and $R^2 = 0.3$, what is the F-statistic for the test that the slope is equal to zero?

Solution: The degrees of freedom are $df_R = 1$ and $df_E = n - 2 = 26$. The F-statistic is

$$F = \frac{MSM}{MSE} = \frac{SSM/df_M}{SSE/df_E} = \frac{SSM}{SSE}\frac{df_E}{df_M} = \frac{R^2}{(1-R^2)}\frac{26}{1} = \frac{0.3 \times 26}{0.7} = 11.14.$$

The $\alpha = 0.05$ critical value for $F_{1,24}$ is 4.26 (from page 1322), and the p-value using that *df* is 0.0027 . The results for $F_{1,26}$ will be similar. We reject $H_0 : \beta_1 = 0$ and conclude that the slope it not zero.

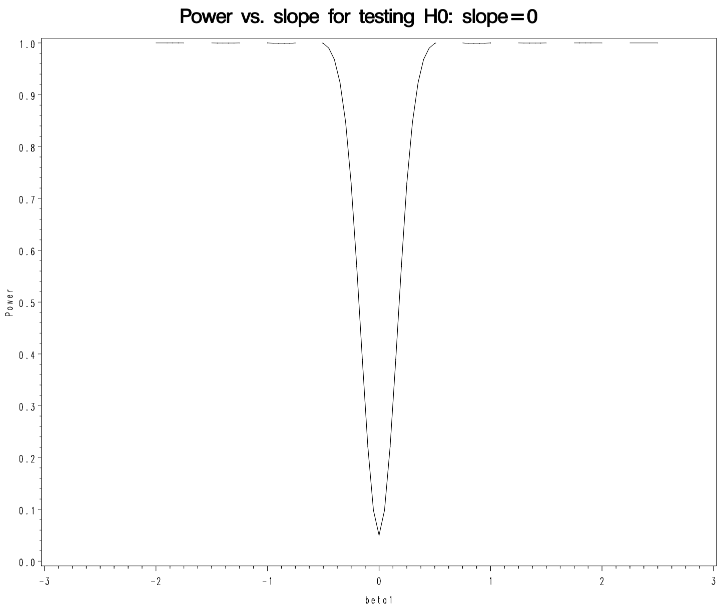

Figure 4: Graph for Problem 8b