**Stat 401 A: Lab 9 self-assessment**

Continuing lab 8’s evaluation of auction prices for antique grandfather clocks. The data in GFCLOCKS.txt include the age of the clock and its price at auction. Consider predicting the price (in British pounds, GBP) from the age (in years). You can ignore the other two variables in the data set. Fit the linear regression line.

1. Predict the price for a 150 year old clock and a 190 year old clock.

2. Report 95% confidence intervals for the average price at auction for a 150 year old clock and for a 190 year old clock.

3. Explain why the confidence interval for a 190 year old clock is almost twice as wide as that for a 150 year old clock.

4. Report 95% prediction intervals for a 150 year old clock and for a 190 year old clock.

5. A 150 year old clock recently sold for 1000 GBP. Do you suspect something unusual about this clock? Briefly explain why or why not?

6. Test for lack of fit to the simple linear regression of price on age. Report the p value and a one-sentence conclusion.

**Answers:**

1. Without rounding: 1379 GBP, 1799 GBP. Better to round (Kelley’s rule) to: 1380 GBP and 1800 GBP

2. After rounding: (1280, 1480) and (1610, 1990) GBP.

3. The prediction at 190 years is further from the mean age.

Note: The ci widths are 200 and 380. The mean age is 145 years; 190 years is very close to the largest age in the data set (194 years).

4. After rounding: (810, 1950) and (1210, 2390) GBP

5. No. The auction price is within the 95% prediction interval.

Note: This question indirectly gets whether a ci or pi is more appropriate. The scenario concerns an individual clock, so the pi is the appropriate comparison distribution.

6. p = 0.74. No evidence of lack of fit.

Note: It is rare to get replicate observations at the same X values in an observational study. Here there are 7 pairs of clocks and one trio with the same age. That gives a 9 df estimate of pure error (variability in price among clocks of the same age).