Paper STATS 330 | **2001 Term Test Solutions** | **Date: 18 September 2001**

1.
   - ϵ's come from a Normal distribution - check using a Normal plot of residuals.
   - errors have constant variance - check using a plot of residuals vs fitted values or squared residuals versus fitted values.
   - errors are independent - check using plot of residuals versus lagged residuals or residuals versus run order.

2. (a) Random scatter (no clear trend)

   (b) A linear trend

   (c) A non-linear (curved) trend

3. (a) A substantial change in the variances and/or covariances of the fitted coefficients would occur.

   (b) A large overall change in the fitted coefficients would occur.

   (c) A large change in the fitted value for that observation would occur.

4. (a) Multicollinearity means that there is a near linear relationship among the the explanatory variables.

   (b) Variance inflation factors (VIF's) are used to detect multicollinearity. A VIF is calculated for each explanatory variable. The VIF is larger for regressors that are involved in a near linear relationship. A VIF of 1 indicates that variable is not related (linearly) to the other explanatory variables. VIF's from 5 - 10 indicate moderate multicollinearity, and VIF's > 10 indicate severe multicollinearity.

   (c) Any one of the following problems:
   - The fitted surface is not stable (small changes in data can cause large changes in fitted coefficients).
   - The standard errors of estimated coefficients are inflated (this makes confidence intervals and hypothesis tests lest precise).
   - Makes model selection more difficult (there may be several subset models that all work approximately as well as each other).

5. (a) The fitted model is:

$$\text{E}(\mathsf{Price}) = -1336.7 + 12.736 \times \mathsf{Age} + 85.815 \times \mathsf{Bidders}$$

   This model suggests that for a fixed number of bidders, the average selling price of a clock increases by 12.736 pounds for each increase of 1 year in the age of the clock. For a clock of a given age, the average selling price of a clock increases by 85.815 pounds for each additional bidder.

   (b) The overall F-test tests the hypothesis H_o:$\beta_{\text{Age}} = \beta_{\text{Bidders}} = 0$. The small p-value gives strong evidence that at least one of the coefficients is not zero. This means the model has some predictive power.

(c) Test the hypothesis $H_o: \beta_{\text{Bidders}} = 100$.
The test statistic is:

$$t_o = \frac{\hat{\beta}_{\text{Bidders}} - 100}{\text{se}(\hat{\beta}_{\text{Bidders}})} = \frac{85.8151 - 100}{8.7058} = -1.629$$

The p-value is calculated as follows:

$$\text{p-value} = 2 \times \Pr(t_{29} \geq 1.629)$$

where t_{29} represents an observation from a t-distribution with 29 df.

6. (a) "Age and Bidders interact" means that the effect that Age has on Price depends on the level of Bidders and vice versa. In each panel of the coplot we see an approximately linear relation between Price and Age but the slopes are different. This indicates that the effect Age has on Price depends on the level of Bidders. If there was no interaction between Age and Bidders each panel should have (approximately) the same slope.

(b) The fitted model is:

$$\text{E}(\textsf{Price}) = 322.8 + 0.8733 \times \textsf{Age} - 93.41 \times \textsf{Bidders} + 1.298 \times \textsf{Age} \times \textsf{Bidders}.$$

To explore the effect Bidders has on Price, fix Age $= c$ and rewrite the fitted model:

$$\text{E}(\textsf{Price}) = 322.8 + 0.8733 \times c + (-93.41 + 1.298 \times c) \times \textsf{Bidders}.$$

Thus if Age is fixed at c, for each additional bidder that participates at the auction the predicted selling price increases by $(-93.41 + 1.298 \times c)$ pounds. For this data Age ranges from 108 to 194 years and thus the increase in Price for each additional bidder ranges from 46.7 pounds for Age $= 108$ to 158.4 pounds for Age $= 194$.