Mth256/656 Test 3 (Take Home) Fall 2011 WITH ANSWER KEY, BELOW

Justify your answers!

1. A probabilistic linear relationship is proposed between hours (x) of therapy and improvement (y) in psychotic patients.
   (a) For a group of 80 patients, present this model being sure to designate both its deterministic and random components.
   (b) Show how the corresponding Z matrix of multipliers of β's would look.
   (c) How would one determine a 90% confidence interval for β_1 ?
   (d) How would one estimate the typical amount that a future observed y value would be off from its (linearly) estimated value?

2. Now suppose that it was also felt that gender (M vs F) could additionally influence improvement with perhaps one gender obtaining more benefit from an additional hour of therapy than did the other gender.
   a. Construct this model and discuss the relevance/meaning of its different components.
   b. Demonstrate how could one test if there is any utility to including anything having to do with gender in the model?

3. Suppose that during June 2008, that (from store to store) Blockbuster store's sales were quadratically associated with numbers of people who entered the stores during June 2008. Propose a regression model and hypothesize as to the sign(s) of the β(s) beyond β_0. Be sure to justify your answer.

1 (a) FOR $i=1,2,\dots,80$; $Y_i = \underbrace{\beta_0+\beta_1\cdot X_i}_{\substack{\text{deterministic}\\ E(Y_i)}} + \underbrace{\epsilon_i}_{\text{random}}$. NOTE: $\boxed{\hat{Y}_i=\hat{\beta}_0+\hat{\beta}_1 X_i}$

(b) $Z = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ \vdots & \vdots \\ 1 & X_{80} \end{pmatrix}$

(c) $\hat{\beta}_1 \pm t_{\frac{1-.9}{2}, df=79} \cdot S_{\hat{\beta}_1} \longleftrightarrow \hat{\beta}_1 \pm 1.645 \frac{s}{\sqrt{SS_{xx}}}$ where $s=\sqrt{\frac{\Sigma(Y_i-\hat{Y}_i)^2}{80-2}}$ and $SS_{xx}=\sum_{i=1}^{80}(x_i-\bar{x})^2$

$t_{\frac{1-.9}{2}, df=79} \approx Z_{\frac{1-.9}{2}} = 1.645$; $S_{\hat{\beta}_1} = \frac{S_{regr}}{\sqrt{SS_{xx}}}$

(d) Estimated standard error of prediction $= s\sqrt{1+\frac{1}{80}+\frac{(X_g-\bar{X})^2}{SS_{xx}}}$ in Section 11.6

2 (a) Think of one gender as the "base", then create a dummy variable for other, e.g. Females $\Rightarrow Z_{2i} = \begin{cases} 1 & \text{if } i \text{ is F} \\ 0 & \text{otherwise} \end{cases}$

Then * $Y_i = \beta_0+\beta_1 X_i + \underbrace{\beta_2 Z_{2i}}_{\text{Avg. effect of F's above M's}} + \underbrace{\beta_3 X_i Z_{2i}}_{\text{interaction between gender \& hours}} + \underbrace{\epsilon_i}_{\text{error}}$

2 (b) H_0: no utility in using Gender; H_0: $0=\beta_2=\beta_3$ Reduced Model
H_A: Not H_0. NOTE: Complete (C) Model is shown in *, above
Reduced (R) Model is $Y_i=\beta_0+\beta_1 X_i+\epsilon_i$; $\hat{Y}_{iR}=\hat{\beta}_{0R}+\hat{\beta}_{1R}X_i$

Test statistic: $\boxed{\dfrac{\frac{SSE_R-SSE_C}{3-1}}{\frac{SSE_C}{n-(3+1)}}} \sim F_{\substack{numdf=2\\ dendf=n-(3+1)}}$; reject H_0 iff data's value of test stat is larger than F_α (under above pair of df's)

Note: $SSE_R=\Sigma(Y_i-\hat{Y}_{iR})^2$

$SSE_C=\Sigma(Y_i-\hat{Y}_{iC})^2$

3 (a) $Y_i = \underbrace{\beta_0+\beta_1 X_i+\beta_2 X_i^2}_{E(Y_i)}+\epsilon_i$, where X_i is # people entered at i

(b) Note that $E(Y_i)$ should be an increasing function of Y_i whose 2^{nd} derivative is probably negative (due to store crowding & less personal attention), which means that β_2 will be negative.