HW3

1\. (a) (i) the function is ~~not~~ linear

(ii) there is no underfitting

(iii) $\beta_0 = 1 \quad \beta_1 = 2 \quad \beta_2 = 1$

(b) (i) the function is not linear

(ii) there is no underfitting

(iii) $a_0 = -3,\ a_1 = 2,\ b_0 = 2,\ b_1 = -1$

(c) (i) the model function is linear

(ii) there is underfitting in the model.

(d) (i) the model is linear

(ii) there is no underfitting

(iii) $\beta_1 = 3 \quad \beta_2 = 3 \quad \beta_3 = -3$

2\. based on the one standard error rule:

$$d_{min} = \arg\min_d \overline{MSE}_d = 3$$

$$\overline{MSE}_{tgt} = \overline{MSE}_{d_{min}} + SE_{d_{min}} = 0.70 + \frac{0.07}{\sqrt{3}} = 0.74$$

Find smallest d such that $\overline{MSE}_d \le \overline{MSE}_{tgt} = 0.74$

when $d = 2$ $\overline{RSS}_d = 0.72 < 0.74$

so $d = 2$ should be selected.

3\. (a) $E\{x^2\} = u^2 + v$

(b) $E\{xy\} = u^2 + v$

(c) $E\{y^2\} = u^2 + v + \sigma^2$

(d) $E\{y|x\} = x$

(e) $E\{y^2|x\} = x^2 + \sigma^2$

4\.

4. (a) $RSS(\beta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - \beta x_i)^2$

$0 = \frac{\partial RSS(\beta)}{\partial \beta}$

$= -2 \sum_i (y_i - \beta x_i) x_i$

$0 / -\frac{1}{2n} = \left[-2 \sum_i (y_i - \beta x_i) x_i \right] / \left(-\frac{1}{2n}\right)$

$= \frac{1}{n} \sum_i (y_i - \beta x_i) x_i$

$= \frac{1}{n} \sum_i y_i x_i - \beta \left(\frac{1}{n} \sum_i x_i^2\right)$

$\hat{\beta} = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$

(b) $\hat{\beta} = \dfrac{\sum_{i=1}^{n} x_i (\beta x_i + \varepsilon_i)}{\sum_{i=1}^{n} x_i^2} = \dfrac{\sum_{i=1}^{n} (\beta x_i^2 + x_i \epsilon_i)}{\sum_{i=1}^{n} x_i^2}$

$= \dfrac{\sum_{i=1}^{n} \beta x_i^2 + \sum_{i=1}^{n} x_i \epsilon_i}{\sum_{i=1}^{n} x_i^2}$

$= \beta + \dfrac{(\frac{1}{n}) \sum_{i=1}^{n} x_i \epsilon_i}{(\frac{1}{n}) \sum_{i=1}^{n} x_i^2}$

(c) $E\{\hat{y} - y | x\} = E\{\hat{\beta} x - (\beta x + \varepsilon) | x\}$

$= E\{\hat{\beta} x | x\} - E\{\beta x | x\} - \cancel{E\{\varepsilon | x\}}^{0}$

$= E\{\hat{\beta} x | x\} - \beta \cdot E\{x | x\}$

$= E\left\{ \left(\beta + \dfrac{\frac{1}{n} \sum_{i=1}^{n} x_i \epsilon_i}{(\frac{1}{n}) \sum_{i=1}^{n} x_i^2} \Bigg|_{x_i = 1 \forall x_i} \right) x \,\Big|\, x \right\} - \beta x$

$= \left(\beta + E\left\{\frac{1}{n} \sum_{i=1}^{n} \epsilon_i \right\}\right) x - \beta x$

$= E\left\{\frac{1}{n} \sum_{i=1}^{n} \epsilon_i \right\} x$

$= 0$ unbiased

(d) $E\{(\hat{y} - y)^2 | x\} = E\{(y - \hat{y})^2 | x\}$

$= E\{(\beta x + \varepsilon - \hat{\beta} x)^2 | x\}$

$= E\{\epsilon^2 + 2\epsilon(\beta x - \hat{\beta} x) + (\beta x - \hat{\beta} x)^2 | x\}$

$= E\{\epsilon^2 | x\} + 2 \cancel{E\{\epsilon | x\}}^{0} E\{(\beta x - \hat{\beta} x) | x\}$

$+ E\{(\beta x - \hat{\beta} x)^2 | x\}$

$= \sigma^2 + E\{(\beta x - \hat{\beta} x)^2 | x\}$

$= \sigma^2 + E\{(\beta x - E\{\hat{\beta} x | x\})^2 | x\} +$

$E\{(E\{\hat{\beta} x | x\} - \hat{\beta} x)^2 | x\} +$

$2(\beta x - E(\hat{\beta} x | x)) \cdot \cancel{E\{(E\{\hat{\beta} x | x\} - \hat{\beta} x) | x\}}^{0}$

$= \sigma^2 + (\cancel{E\{(\hat{y} - y) | x\}}^{0})^2 + E\{(\hat{\beta} x - E\{$

$\hat{\beta} x | x\})^2 | x\}$

$= \sigma^2 + E\{(\hat{\beta} x)^2 | x\} - E\{(\hat{\beta} x) | x\}^2$

$= \sigma^2 + E\{(\hat{\beta} x)^2 | x\} - (\beta x)^2$

$= \sigma^2 + E\{((\beta + \frac{1}{n} \sum_{i=1}^{n} \epsilon_i) x)^2 | x\} - (\beta x)^2$

$= \sigma^2 + \cancel{E\{(\beta x)^2 | x\} - (\beta x)^2}^{0} + E\{(\frac{1}{n} \sum_{i=1}^{n} \epsilon_i)^2$

$x^2 | x\} + \cancel{E\{2\beta (\frac{1}{n}) \sum_{i=1}^{n} \epsilon_i x^2 | x\}}^{0}$

$= \sigma^2 + \frac{\sigma^2}{n} \cdot x^2$