Homework 1.

Question 1.

1. Show that β_0 is unbiased

$$\begin{aligned}
y_i &= \beta_0 + \beta_1 x_i + \epsilon_i \\
\mathrm{E}(\beta_0) &= \mathrm{E}(y_i - \beta_1 x_i - \epsilon_i) \\
\mathrm{E}(\beta_0) &= \mathrm{E}(y_i) - \beta_1 \mathrm{E}(x_i) - \mathrm{E}(\epsilon_i) \\
\mathrm{E}(\beta_0) &= \bar{y} - \beta_1 \bar{x} - 0 \\
\mathrm{E}(\beta_0) &= \bar{y} - \beta_1 \bar{x} \\
\mathrm{E}(\beta_0) &= \beta_0 \\
\beta_0 &= \bar{y} - \beta_1 \bar{x}
\end{aligned}$$

$$\begin{aligned}
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\
\mathrm{E}(\hat{\beta}_0) &= \mathrm{E}(\bar{y} - \hat{\beta}_1 \bar{x}) \\
\mathrm{E}(\hat{\beta}_0) &= \mathrm{E}(\bar{y}) - \mathrm{E}(\hat{\beta}_1 \bar{x}) \\
\mathrm{E}(\hat{\beta}_0) &= \bar{y} - \bar{x}\,\mathrm{E}(\hat{\beta}_1) \\
\mathrm{E}(\hat{\beta}_1) &= \beta_1 && \text{(we showed this in class)} \\
\mathrm{E}(\hat{\beta}_0) &= \bar{y} - \beta_1 \bar{x} \\
\mathrm{E}(\hat{\beta}_0) &= \beta_0
\end{aligned}$$

2. Determine the variance of β_0

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$$

$$\begin{aligned}
\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1\bar{x} && \text{(from class)} \\
\mathrm{Var}(\hat{\beta}_0) &= \mathrm{Var}(\bar{y} - \hat{\beta}_1\bar{x}) \\
\mathrm{Var}(\hat{\beta}_0) &= \mathrm{Var}(\bar{y}) + \mathrm{Var}(\hat{\beta}_1\bar{x}) + 2\,\mathrm{Cov}(\bar{y}, -\hat{\beta}_1\bar{x})
\end{aligned}$$

$$\begin{aligned}
\bar{y} &= \frac{1}{n}\sum_{i=1}^{n} y_i \\
\mathrm{Var}(\bar{y}) &= \mathrm{Var}(\frac{1}{n}\sum_{i=1}^{n} y_i) \\
\mathrm{Var}(\bar{y}) &= (\frac{1}{n})^2\sum_{i=1}^{n} \mathrm{Var}(y_i) \\
\mathrm{Var}(\bar{y}) &= (\frac{1}{n})^2\sum_{i=1}^{n} \mathrm{Var}(\beta_0 + \beta_1 x_i + \epsilon_i) \\
\mathrm{Var}(\bar{y}) &= (\frac{1}{n})^2\sum_{i=1}^{n} \mathrm{Var}(\beta_0) + \mathrm{Var}(\beta_1 x_i) + \mathrm{Var}(\epsilon_i) \\
\mathrm{Var}(\bar{y}) &= (\frac{1}{n})^2\sum_{i=1}^{n} 0 + 0 + \sigma^2 \\
\mathrm{Var}(\bar{y}) &= (\frac{1}{n})^2.(n\sigma^2) \\
\mathrm{Var}(\bar{y}) &= \frac{\sigma^2}{n}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(-\beta_1\bar{x}) &= (-\bar{x})^2\,\mathrm{Var}(\hat{\beta}_1) \\
\mathrm{Var}(-\beta_1\bar{x}) &= (-\bar{x})^2\,\mathrm{Var}(\hat{\beta}_1) \\
\mathrm{Var}(-\beta_1\bar{x}) &= (-\bar{x})^2\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})} && \text{(we showed this in class)}
\end{aligned}$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x}\operatorname{Cov}(\bar{y}, \hat{\beta}_1)$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x}\operatorname{Cov}\left(\frac{1}{n}\sum_{i=1}^{n} y_i, \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) \qquad \text{(definition from class notes)}$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x} \cdot \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\operatorname{Cov}\left(\sum_{i=1}^{n} y_i, \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})\right)$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x} \cdot \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\operatorname{Cov}\left(\sum_{i=1}^{n} y_i, \sum_{i=1}^{n}(x_i - \bar{x})(y_i) - \bar{y}\sum_{i=1}^{n}(x_i - \bar{x})\right)$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x} \cdot \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\operatorname{Cov}\left(\sum_{i=1}^{n} y_i, \sum_{i=1}^{n}(x_i - \bar{x})(y_i)\right) \qquad \text{(special zero from class)}$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = \bar{x} \cdot \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\sum_{i=1}^{n}(x_i - \bar{x})\sum_{i=1}^{n}\operatorname{Cov}(y_i, y_i)$$

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

$$\operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) = 0$$

$$\begin{aligned}\operatorname{Var}(\hat{\beta}_0) &= \operatorname{Var}(\bar{y}) + \operatorname{Var}(\hat{\beta}_1 \bar{x}) + \operatorname{Cov}(\bar{y}, \hat{\beta}_1 \bar{x}) \\ \operatorname{Var}(\hat{\beta}_0) &= \operatorname{Var}(\bar{y}) + \operatorname{Var}(\hat{\beta}_1 \bar{x}) + 0 \\ &\boxed{\operatorname{Var}(\hat{\beta}_0) = \frac{\sigma^2}{n} + \bar{x}^2\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})}}\end{aligned}$$

3(a). Show that $\sum_{i=1}^{n} \hat{y}_i \cdot e_i = 0$

$$\sum_{i=1}^{n} \hat{y}_i \cdot e_i = 0$$

$$\sum_{i=1}^{n} (\hat{\beta}_0 + \hat{\beta}_1 x_i)(e_i) = 0$$

$$\sum_{i=1}^{n} \hat{\beta}_0(e_i) + \hat{\beta}_1(x_i)(e_i) = 0$$

$$\sum_{i=1}^{n} \hat{\beta}_0(e_i) + \sum_{i=1}^{n} \hat{\beta}_1(x_i)(e_i) = 0$$

$$\hat{\beta}_0 \sum_{i=1}^{n}(e_i) + \hat{\beta}_1 \sum_{i=1}^{n}(x_i)(e_i) = 0$$

$$\sum_{i=1}^{n}(e_i) = 0 \qquad \text{(from property 1)}$$

$$\sum_{i=1}^{n}(x_i)(e_i) = 0 \qquad \text{(from property 4)}$$

$$\hat{\beta}_0 \cdot 0 + \hat{\beta}_1 \cdot 0 = 0$$

$$0 = 0$$

3(b). Show that regression line passes through $(\overline{x}, \overline{y})$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

$$\text{let } x_i = \overline{x}$$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \overline{x}$$

$$\hat{y}_i = \overline{y} - \beta_1 \overline{x} + \hat{\beta}_1 \overline{x} \qquad \text{(from normal equations)}$$

$$\hat{y}_i = \overline{y}$$

this shows that the regression line passes through $(\overline{x}, \overline{y})$