Exercise 1.4

(a) Do you think that the value of b depends on the variance of the measurement errors? Why?

$$\beta = \frac{Cov(x,y)}{Var(x)}$$

Variance of the measurement error contribute to the variance of x, hence value of b is dependent on the variance of the measurement error.

(b) Show that $b = \beta + \sum_{i=1}^{n} \frac{(x_i-\bar{x})(\epsilon_i-\bar{\epsilon})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$

$$\begin{aligned}
b &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
y_i - \bar{y} &= (\alpha + \beta x_i + \epsilon_i) - (\alpha + \beta\bar{x} + \bar{\epsilon}) \\
&= \beta(x_i - \bar{x}) - (\epsilon_i - \bar{\epsilon})
\end{aligned}$$

$$\begin{aligned}
b &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})[\beta(x_i-\bar{x}) + (\epsilon_i-\bar{\epsilon})]}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
&= \beta\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{\sum_{i=1}^{n}(x_i-\bar{x})(\epsilon_i-\bar{\epsilon})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\
&= \beta + \frac{\sum_{i=1}^{n}(x_i-\bar{x})(\epsilon_i-\bar{\epsilon})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}
\end{aligned}$$

(c) Show that $\epsilon_i = \epsilon_i^* - \beta v_i$

$$\begin{aligned}
y_i = \alpha + \beta x_i^* + \epsilon_i^* &= \alpha + \beta x_i + \epsilon_i \\
\alpha + \beta x_i^* + \epsilon_i^* &= \alpha + \beta(x_i^* + v_i) + \epsilon_i \\
\alpha + \beta x_i^* + \epsilon_i^* &= \alpha + \beta x_i^* + \beta v_i + \epsilon_i \\
\epsilon_i^* &= \beta v_i + \epsilon_i \\
\epsilon_i &= \epsilon_i^* - \beta v_i
\end{aligned}$$

(d) Show that the covariance between x_i and ϵ_i equal to $-\beta\sigma_v^2$

Notes

$$\sum_{i=1}^{n}(x_i^* - \bar{x}^*)(\epsilon_i^* - \bar{\epsilon}^*) = 0$$
$$\sum_{i=1}^{n}(x_i^* - \bar{x}^*)(v_i - \bar{v}) = 0$$
$$\sum_{i=1}^{n}(v_i - \bar{v})(\epsilon_i^* - \bar{\epsilon}^*) = 0$$

$$\begin{aligned}
Cov(x_i, \epsilon_i) &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\epsilon_i - \bar{\epsilon})}{n-1} \\
&= \frac{\sum_{i=1}^{n}(x_i^* + v_i - \bar{x}^* - \bar{v})(\epsilon_i^* - \beta v_i - \bar{\epsilon}^* + \beta\bar{v})}{n-1} \\
&= \frac{\sum_{i=1}^{n}([x_i^* - \bar{x}^*] + [v_i - \bar{v}])([\epsilon_i^* - \bar{\epsilon}^*] - \beta[v_i - \bar{v}])}{n-1} \\
&= -\frac{\sum_{i=1}^{n}(v_i - \bar{v})\beta(v_i - \bar{v})}{n-1} \\
&= -\beta Var(v_i) \\
&= -\beta\sigma_v^2
\end{aligned}$$

(e) Show that for large sample size n we get $b - \beta = \frac{-\beta\sigma^2}{\sigma_*^2 + \sigma_v^2}$

$$\begin{aligned}
b - \beta &= \frac{Cov(x_i, \epsilon_i)}{Var(x_i)} \\
&= \frac{-\beta\sigma_v^2}{Var(x_i^* + v_i)} \\
&= \frac{-\beta\sigma_v^2}{Var(x_i^*) + Var(v_i)} \\
&= \frac{-\beta\sigma_v^2}{\sigma_*^2 + \sigma_v^2}
\end{aligned}$$

(f) Compute the approximate bias $(b - \beta)$ for $\beta = 1$ in the cases SN = 1, SN = 3, SN = 10

$$\begin{aligned} b - \beta &= \frac{-\beta\sigma_v^2}{\sigma_*^2 + \sigma_v^2} \\ &= \frac{-\sigma_v^2}{\sigma_*^2 + \sigma_v^2} \\ &= -\frac{1}{\frac{\sigma_*^2}{\sigma_v^2} + 1} \\ &= -\frac{1}{SN + 1} \end{aligned}$$

$$\begin{aligned} SN &= 1 & b - \beta &= -\frac{1}{2} \\ SN &= 3 & b - \beta &= -\frac{1}{4} \\ SN &= 10 & b - \beta &= -\frac{1}{11} \end{aligned}$$