Homework 4 - Answers

7 (i) $E(u|inc) = E(\sqrt{inc}\, e \,|\, inc) = \sqrt{inc}\, E(e|inc)$. But since e and inc are independent $E(e|inc) = E(e) = 0$. Hence, $E(u|inc) = 0$.

(ii) $V(u|inc) = V(\sqrt{inc}\, e \,|\, inc) = inc\, V(e|inc) = inc\, \sigma_e^2$ since inc and e are independent

$V(sav|inc) = V(u|inc) = inc\, \sigma_e^2$. Since $\sigma_e^2 > 0$, $V(sav|inc)$ increases with inc.

(iii) lower income levels are associated with a higher fraction of income devoted to consumption. Hence, smaller savings and smaller variation of savings.

10 (i) From class notes,

$$\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})(u_i - \bar{u})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \bar{x})u_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \beta_1 + \sum_{i=1}^{n} w_i u_i$$

where $w_i = d_i / SST_x$, with $d_i = (x_i - \bar{x})$, $SST_x = \sum_{i=1}^{n}(x_i - \bar{x})^2$

(ii) $$E[(\hat{\beta}_1 - \beta_1)\bar{u}] = E\left(\left[\sum_{i=1}^{n} w_i u_i\right]\bar{u}\right) = E\left(\sum_{i=1}^{n} w_i u_i \bar{u}\right)$$

$$= \sum_{i=1}^{n} E(w_i u_i \bar{u})$$

But
$$E(w_i u_i \bar{u}) = E(E(w_i u_i \bar{u} \,|\, x_1 \cdots x_n))$$
$$= E(w_i\, E(u_i \bar{u} \,|\, x_1 \cdots x_n))$$
$$= E\left(w_i \frac{1}{n}\sigma^2\right)$$
$$= \frac{1}{n}\sigma^2 E(w_i)$$

$$E((\hat{\beta}_1 - \beta_1)\bar{u}) = \sum_{i=1}^{n} E(w_i u_i \bar{u}) = \frac{1}{n}\sigma^2 \sum_{i=1}^{n} E(w_i)$$
$$= \frac{\sigma^2}{n} E\left(\sum_{i=1}^{n} w_i\right) = \frac{\sigma^2}{n} E\left(\frac{1}{SST_x}\sum_{i=1}^{n} d_i\right)$$

$$= 0$$

(iii) From class

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$
$$= \beta_0 + \beta_1 \bar{X} + \bar{U} - \hat{\beta}_1 \bar{X}$$
$$= \beta_0 - (\hat{\beta}_1 - \beta_1)\bar{X} + \bar{U}$$

(iv) $$V(\hat{\beta}_0 | X_1 \dots X_n) = E\left([\hat{\beta}_0 - \beta_0]^2 | X_1 \dots X_n\right)$$
$$= E\left([\bar{U} - (\hat{\beta}_1 - \beta_1)\bar{X}]^2 | X_1 \dots X_n\right)$$
$$= E(\bar{U}^2 | X_1, \dots, X_n) + \bar{X}^2 E((\hat{\beta} - \beta_1)^2 | X_1 \dots X_n).$$
$$- 2E(\bar{U}(\hat{\beta}_1 - \beta_1)\bar{X} | X_1 \dots X_n)$$

By (ii), the last term is zero.

$$E(\bar{U}^2 | X_1, \dots, X_n) = E\left(\frac{1}{n^2}\left[\sum_{i=1}^n U_i\right]^2 | X_1 \dots X_n\right).$$
$$= \frac{1}{n^2}\left\{E\left(\sum_{i=1}^n U_i^2 + \sum_{i \neq j} U_i U_j \,\Big|\, X_1 \dots X_n\right)\right\}.$$
$$= \frac{1}{n^2}\left\{n\sigma^2 + 0\right\} = \sigma^2/n$$

since $E(U_i^2 | X_1 \dots X_n) = \sigma^2$ and $\{U_i\}_{i=1}^n$ forms an independent sequence.

Then,

$$V(\hat{\beta}_0 | X_1 \dots X_n) = \frac{\sigma^2}{n} + \frac{\bar{X}^2 \sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sigma^2}{n} + \frac{\bar{X}^2 \sigma^2}{SST_x}$$

where, $V(\hat{\beta}_1 | X_1 \dots X_n) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$ from class notes

(v) $$V(\hat{\beta}_0 | X_1 \dots X_n) = \frac{\sigma^2 SST_x + \bar{X}^2 \sigma^2 n}{n\, SST_x} = \frac{n\sigma^2\left(\frac{SST_x}{n} + \bar{X}^2\right)}{n\, SST_x}$$
$$= \frac{n\sigma^2 \frac{1}{n}\sum_{i=1}^n x_i^2}{n\, SST_x} \quad \text{from the hint.}$$
$$= \frac{\sigma^2 \frac{1}{n}\sum x_i^2}{n SST_x} \quad \text{as in [2.58]}$$

11 (i) Divide a group of students that are similar in several observable variables, for example, same GPA in high school, same classes taken, same education of parents, etc., in sub groups. Assign to each subgroup $0, 1, 2, \ldots, N$ hours of SAT preparation course and observe $\underline{sat}$. An average score for each group can be calculated and compared.

(ii) High school GPA, education level of parents, student took calculus in high school.

(iii) $\beta_1 > 0$

(iv) $E(sat \mid hours) = \beta_0 + \beta_1 hours$.

$= \beta_0$ if $hours = 0$

Hence, β_0 is the expected sat score given no hours on course preparation.

13 (i) $\sum_{i=1}^{n} x_i = \sum_{x_i=0} x_i + \sum_{x_i=1} x_i = 0 + n_1$. But since $n = n_0 + n_1$,

$\sum_{i=1}^{n} 1 = n_0 + \sum_{i=1}^{n} x_i$. Then, $\sum_{i=1}^{n} (1 - x_i) = n_0$.

$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{n_1}{n}$ and $1 - \bar{x} = 1 - \frac{n_1}{n} = \frac{n - n_1}{n} = \frac{n_0}{n}$

$\bar{x}$ is the proportion of 1's in $\{x_i\}_{i=1}^{n}$.

(ii) Note that $y_i = (1 - x_i) y_i + y_i x_i$. Hence,

$$\bar{y}_0 = \frac{1}{n_0} \sum_{i=1}^{n} (1 - x_i) y_i \quad \text{and} \quad \bar{y}_1 = \frac{1}{n_1} \sum_{i=1}^{n} y_i x_i$$

(iii)
$$\frac{1}{n} \sum_{i=1}^{n} y_i = \frac{1}{n} \sum_{i=1}^{n} (1 - x_i) y_i + \frac{1}{n} \sum_{i=1}^{n} x_i y_i$$

$$= \frac{n_0}{n} \frac{1}{n_0} \sum_{i=1}^{n} y_i (1 - x_i) + \frac{n_1}{n} \frac{1}{n_1} \sum_{i=1}^{n} x_i y_i$$

$$= \frac{n_0}{n} \bar{y}_0 + \frac{n_1}{n} \bar{y}_1 = (1 - \bar{x}) \bar{y}_0 + \bar{x} \bar{y}_1$$

(iv) $\frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$. Hence, $\frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\bar{x})^2 = \bar{x} - (\bar{x})^2$

$= \bar{x}(1-\bar{x})$.

(v) From (ii) $1/n_1 \sum_{i=1}^{n} x_i y_i = \bar{y}_1 \Leftrightarrow (n/n_1)\, 1/n \sum_{i=1}^{n} x_i y_i = \bar{y}_1 \Leftrightarrow$
$1/n \sum_{i=1}^{n} x_i y_i = (n_1/n)\,\bar{y}_1 = \bar{x}\,\bar{y}_1$ from (i).

From (iii) $\bar{x}\,\bar{y} = \bar{x}\left[(1-\bar{x})\,\bar{y}_0 + \bar{x}\,\bar{y}_1\right]$
$= \bar{x}\left[\bar{y}_0 + \bar{x}(\bar{y}_1 - \bar{y}_0)\right]$.

Hence,
$$1/n \sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} = \bar{x}\bar{y}_1 - \bar{x}\bar{y}_0 - \bar{x}^2(\bar{y}_1 - \bar{y}_0)$$
$$= (\bar{y}_1 - \bar{y}_0)(\bar{x} - \bar{x}^2) = (\bar{y}_1 - \bar{y}_0)(1-\bar{x})\,\bar{x}$$

(vi) $$\hat{\beta}_1 = \frac{1/n \sum x_i y_i - \bar{x}\bar{y}}{\frac{1}{n}\sum x_i^2 - \bar{x}^2} = \frac{(\bar{y}_1 - \bar{y}_0)(1-\bar{x})\,\bar{x}}{\bar{x}(1-\bar{x})} = \bar{y}_1 - \bar{y}_0$$

$$\hat{\beta}_0 = \bar{y} - \bar{x}\hat{\beta}_1 = (1-\bar{x})\bar{y}_0 + \bar{x}\bar{y}_1 - \bar{x}\bar{y}_1 + \bar{x}\bar{y}_0$$
$$= \bar{y}_0.$$