HW 13 - Solutions

Exercise 1:

1) $Y_i = \underbrace{\alpha + \beta x_i}_{\substack{\text{constant} \\ \text{(non random)}}} + U_i$ $\leftarrow N(0, \sigma^2)$

so Y_i is Gaussian with mean $E[Y_i] = \alpha + \beta x_i + 0$
and variance $Var(Y_i) = Var(U_i) = \sigma^2$

2) $$L = \prod_{i=1}^{n} \frac{e^{-\frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$ (joint p.d.f given data)

$$\ell = -\frac{n}{2}\ln(\sigma^2) - n\ln(\sqrt{2\pi}) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \alpha - \beta x_i)^2$$

3) $$\frac{\partial \ell}{\partial \alpha} = \frac{-1}{\sigma^2}\sum_{i=1}^{n}(\alpha + \beta x_i - y_i) = 0 \iff \alpha + \beta \overline{x_n} = \overline{y_n}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{-1}{\sigma^2}\sum_{i=1}^{n} x_i(\alpha + \beta x_i - y_i) = 0$$

$$\iff \sum_{i=1}^{n} x_i(\beta(x_i - \overline{x_n}) + \overline{y_n} - y_i) = 0$$

$$\iff \beta \sum_i (x_i - \overline{x_n})^2 = \sum_i (x_i - \overline{x_n})(y_i - \overline{y_n})$$

$$\frac{\partial \ell}{\partial \sigma^2} = \frac{-n}{\cancel{2}\sigma^2} + \frac{1}{\cancel{2}\sigma^4} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 = 0$$

$$\Longleftrightarrow \left[\begin{array}{l} \hat{\sigma}^2 = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta} x_i)^2 \\[2ex] \hat{\alpha} = \overline{y_n} - \hat{\beta}\, \overline{x_n} \\[2ex] \hat{\beta} = \dfrac{\sum_i (x_i - \overline{x_n})(y_i - \overline{y_n})}{\sum_i (x_i - \overline{x_n})^2} \end{array}\right.$$

Exercise 2:

① $$E[\hat{\beta}] = \frac{\sum (x_i - \overline{x_n}) E[Y_i - \overline{Y_n}]}{\sum (x_i - \overline{x_n})^2}$$

$Y_i \sim N(\alpha + \beta x_i, \sigma^2)$ so $E[Y_i] = \alpha + \beta x_i$

and $$E[\overline{Y_n}] = \frac{(\alpha+\beta x_1) + \dots + (\alpha + \beta x_n)}{n}$$

$$= \alpha + \beta \overline{x_n}$$

and $$E[\hat{\beta}] = \frac{\sum (x_i - \overline{x_n}) E[Y_i - \overline{Y_n}]}{\sum (x_i - \overline{x_n})^2}$$

$$= \frac{\sum (x_i - \overline{x_n}) \beta (x_i - \overline{x_n})}{\sum (x_i - \overline{x_n})^2}$$

$$= \beta$$

$$\hat{\alpha} = \overline{Y_n} - \hat{\beta}\,\overline{x_n} \qquad E[\hat{\alpha}] = (\alpha + \beta\overline{x_n}) - \beta\overline{x_n}$$

$$= \alpha$$

② $$\sum_i (x_i - \overline{x_n})(Y_i - \overline{Y_n}) \qquad = n\overline{x_n} - n\overline{x_n} = 0$$

$$= \sum_i (x_i - \overline{x_n})Y_i - \overline{Y_n}\sum_i (x_i - \overline{x_n})$$

Since $(Y_1, \dots, Y_n)$ are independent

$$Var\left(\sum_i (x_i - \overline{x_n})Y_i\right) = \sum_i (x_i - \overline{x_n})^2 \sigma^2$$

$$Var(\hat{\beta}) = \frac{\sum_i (x_i - \overline{x_n})^2 \sigma^2}{\left(\sum_i (x_i - \overline{x_n})^2\right)^2} = \frac{\sigma^2}{\sum_i (x_i - \overline{x_n})^2}$$

③ $$\mathrm{Cov}(\hat{\beta}, \bar{Y}_n) = \frac{1}{\sum (x_i - \bar{x}_n)^2} \sum_i (x_i - \bar{x}_n) \underbrace{\mathrm{Cov}(Y_i, \bar{Y}_n)}_{= \frac{\sigma^2}{n}}$$

$$= \frac{\sigma^2}{n \sum (x_i - \bar{x}_n)^2} \underbrace{\sum_{i=1}^{n} (x_i - \bar{x}_n)}_{= 0}$$

$$= 0$$

$$\mathrm{Var}(\hat{\alpha}) = \mathrm{Var}(\bar{Y}_n - \hat{\beta}\,\bar{x}_n)$$

$$= \mathrm{Var}(\bar{Y}_n) - 2\bar{x}_n \mathrm{Cov}(\hat{\beta}, \bar{Y}_n) + \bar{x}_n^2 \mathrm{Var}(\hat{\beta})$$

$$= \frac{\sigma^2}{n} - 0 + \bar{x}_n^2 \frac{\sigma^2}{\sum (x_i - \bar{x}_n)^2}$$

④ $$\hat{\alpha} \sim N\left(\alpha,\ \sigma \sqrt{\frac{1}{n} + \frac{\bar{x}_n^2}{\sum (x_i - \bar{x}_n)^2}}\right)$$

$$\hat{\beta} \sim N\left(\beta,\ \frac{\sigma}{\sqrt{\sum (x_i - \bar{x}_n)^2}}\right)$$

Exercise 3: Part (i)

* $e_i = Y_i - \hat{Y}_i = (\alpha + \beta x_i) + U_i - (\hat{\alpha} + \hat{\beta} x_i)$

$$E[e_i] = \alpha + \beta x_i + 0 - \alpha - \beta x_i = 0$$

* $$Cov(Y_i, \hat{Y}_i) = Cov(U_i, \hat{\alpha} + \hat{\beta} x_i)$$

$$= Cov(U_i, \bar{Y}_n - \hat{\beta}\bar{x}_n + \hat{\beta} x_i)$$

$$= Cov(U_i, \bar{Y}_n) + (x_i - \bar{x}_n) Cov(U_i, \hat{\beta})$$

$$= \frac{\sigma^2}{n} + \frac{(x_i - \bar{x}_n)^2}{\sum (x_i - \bar{x}_n)^2} \sigma^2$$

* $$Var(\hat{Y}_i) = Var(\hat{\alpha} + \hat{\beta} x_i)$$

$$= Var(\bar{Y}_n + \hat{\beta}(x_i - \bar{x}_n))$$

$Cov(\bar{Y}_n, \hat{\beta}) = 0$

$$= Var(\bar{Y}_n) + (x_i - \bar{x}_n)^2 Var\,\hat{\beta}$$

$$= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x}_n)^2}{\sum_i (x_i - \bar{x}_n)^2} \right)$$

$* \, Var(Y_i - \hat{Y}_i) = Var\, Y_i - 2\, Cov(Y_i, \hat{Y}_i) + Var(\hat{Y}_i)$

$$= \sigma^2 \left(1 - \frac{1}{n} - \frac{(x_i - \bar{x}_n)^2}{\sum_i (x_i - \bar{x}_n)^2} \right)$$

Part (ii)

$* \, E[e_p] = E[\alpha + \beta x + U - (\hat{\alpha} + \hat{\beta} x)]$

$$= \alpha + \beta x + 0 - \alpha - \beta x = 0$$

$*$ Y is independent of $\hat{Y}$ so

$Var\, e_p = Var\, Y + Var\, \hat{Y}$

$$= \sigma^2 + Var(\bar{Y}_n + \hat{\beta}(x - \bar{x}_n))$$

$$= \sigma^2 \left(1 + \frac{1}{n} + \frac{(x - \bar{x}_n)^2}{\sum_i (x_i - \bar{x}_n)^2} \right)$$