1.) $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

a.) $\mathbb{E}[\hat{Y}_h \mid X = X_h] = \mathbb{E}[\hat{\beta}_0 + \hat{\beta}_1 X_h + \varepsilon_h]$

$= \beta_0 + \beta_1 X_h$, since $\hat{\beta}_0, \hat{\beta}_1$ are unbiased and $\mathbb{E}[\varepsilon] = 0$

b.) $Var(\hat{\beta}_0) = \dfrac{\sigma^2 \Sigma X_i^2}{n \Sigma (X_i - \bar{X})^2}$

$Var(\hat{\beta}_1) = \dfrac{\sigma^2}{\Sigma (X_i - \bar{X})^2}$

$Cov(\bar{Y}, \hat{\beta}_1) = 0$

c.) $Var(\hat{Y}_h) = Var(\hat{\beta}_0 + \hat{\beta}_1 X_h)$

$= Var(\bar{Y} - \hat{\beta}_1 \bar{X} + \hat{\beta}_1 X_h)$

$= Var(\bar{Y}) + Var(\hat{\beta}_1 (X_h - \bar{X}))$, since $Cov(\bar{Y}, \hat{\beta}_1) = 0$

$= \dfrac{\sigma^2}{n} + \sigma^2 \left[\dfrac{(X_h - \bar{X})^2}{\Sigma (X_i - \bar{X})^2} \right]$

$= \sigma^2 \left[\dfrac{1}{n} + \dfrac{(X_h - \bar{X})^2}{\Sigma (X_i - \bar{X})^2} \right]$

d.) $\hat{Y}_h \sim N(\beta_0 + \beta_1 X_h, Var\{\hat{Y}_h\})$

e.) As $|X_h - \bar{X}|$ increases, $Var(\hat{Y}_h)$ increases due to the term $(X_h - \bar{X})^2$ in the expression $Var(\hat{Y}_h)$

(b) f.) $\hat{Y}_h - \mathbb{E}[Y|X=X_h] \sim N(0, Var(\hat{Y}_h))$

$$\frac{\hat{Y}_h - \mathbb{E}[Y|X=X_h]}{\sqrt{\sigma^2\left[\frac{1}{n} + \sum_i \frac{(X_h-\bar{X})^2}{(X_i-\bar{X})^2}\right]}} \sim N(0,1)$$

$$\frac{\hat{Y}_h - \mathbb{E}[Y|X=X_h]}{\sqrt{s^2\left[\frac{1}{n} + \frac{(X_h-\bar{X})^2}{\Sigma(X_i-\bar{X})^2}\right]}} \sim t_{n-2}$$

g.) $(1-\alpha)100\%$ C.I. for $\mu(X_h)$ is:

$$\hat{\mu}(X_h) \pm s_{\hat{Y}_h} \cdot t_{n-2}(1-\alpha/2)$$

where $s_{\hat{Y}_h} = \sqrt{s^2\left[\frac{1}{n} + \frac{(X_h-\bar{X})^2}{\Sigma(X_i-\bar{X})^2}\right]}$

h.) Prediction interval is

$$\hat{\mu}(X_h) \pm t_{n-2}(1-\alpha/2) \cdot \sqrt{s^2\left[1+\frac{1}{n} + \frac{(X_h-\bar{X})^2}{\Sigma(X_i-\bar{X})^2}\right]}$$

i.) The conf. int. margin depends on the variance of the mean est. $\hat{\mu}(X_h)$, whereas the prediction interval margin incorporates the variance of the mean est as well as the variance of an individual observation, hence the additional +1 in the ~~variance~~ standard error of the P.I.

(3) $Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_{p-1} X_{p-1,i} + \varepsilon_i$

a.) $\beta_1 = E[Y \mid X_1 = x+1] - E[Y \mid X_1 = x]$, where all other X_j are equal

That is, β_1 is the difference in mean response comparing two subpopulations that differ in X_1 by 1 unit, and with all other covariates equal.

b.) $Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ → response vector $(n \times 1)$

$X = \begin{pmatrix} 1 & X_{11} & \cdots & X_{1,p-1} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & & X_{n,p-1} \end{pmatrix}$ → design matrix $(n \times p)$

$\beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{p-1} \end{pmatrix}$ → coeff. vector $(p \times 1)$

$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$ → error vector $(n \times 1)$

③ c.) $\mathcal{L}(\underline{\beta}, \sigma^2 | X, Y) = (2\pi\sigma^2)^{-n/2} \exp\left\{\frac{-1}{2\sigma^2}(Y - X\beta)^T(Y - X\beta)\right\}$

$$\ell(\underline{\beta}, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(Y - X\beta)^T(Y - X\beta)$$

$$\frac{\partial \ell}{\partial \beta} = \frac{1}{\sigma^2} X^T(Y - X\beta)$$

d.) $\frac{\partial \ell}{\partial \beta} = 0 \Rightarrow X^T(Y - X\hat{\beta}) = 0$

$$X^T Y - X^T X\hat{\beta} = 0$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

e.) $\mathrm{Var}(\hat{\beta}) = \mathrm{Var}((X^T X)^{-1} X^T Y) = (X^T X)^{-1} X^T \mathrm{Var}(Y) X (X^T X)^{-1}$

$$= (X^T X)^{-1} X^T [\sigma^2 \cdot I_n] X (X^T X)^{-1}$$

$$= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1}$$

$$= \sigma^2 (X^T X)^{-1}$$

3\. f) $\begin{pmatrix}\hat\beta_0\\ \hat\beta_1\end{pmatrix} = (X^TX)^{-1}X^TY$

$$X^TX = \begin{pmatrix}1 & \cdots & 1\\ x_1 & \cdots & x_n\end{pmatrix}\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix} = \begin{pmatrix}n & \Sigma X_i\\ \Sigma X_i & \Sigma X_i^2\end{pmatrix}$$

$$(X^TX)^{-1} = \frac{1}{n\Sigma X_i^2 - (\Sigma X_i)^2}\begin{pmatrix}\Sigma X_i^2 & -\Sigma X_i\\ -\Sigma X_i & n\end{pmatrix}$$

$$X^TY = \begin{pmatrix}1 & \cdots & 1\\ x_1 & \cdots & x_n\end{pmatrix}\begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix} = \begin{pmatrix}\Sigma Y_i\\ \Sigma X_iY_i\end{pmatrix}$$

$$(X^TX)^{-1}X^TY = \frac{1}{n\Sigma X_i^2 - (\Sigma X_i)^2}\begin{pmatrix}\Sigma X_i^2 & -\Sigma X_i\\ -\Sigma X_i & n\end{pmatrix}\begin{pmatrix}\Sigma Y_i\\ \Sigma X_iY_i\end{pmatrix}$$

$$= \frac{1}{n\Sigma X_i^2 - (\Sigma X_i)^2}\begin{pmatrix}\Sigma X_i^2 \cdot \Sigma Y_i - \Sigma X_i \Sigma X_iY_i\\ n\Sigma X_iY_i - n^2\bar X\bar Y\end{pmatrix}$$

$$= \frac{1}{n\Sigma X_i^2 - n^2\bar X^2}\begin{pmatrix}n\bar Y\Sigma X_i^2 - n\bar X\Sigma X_iY_i\\ n\Sigma X_iY_i - n^2\bar X\bar Y\end{pmatrix}$$

$$= \frac{1}{\Sigma X_i^2 - n\bar X^2}\begin{pmatrix}\bar Y\Sigma X_i^2 - \bar X\Sigma X_iY_i\\ \Sigma X_iY_i - n\bar X\bar Y\end{pmatrix}$$

$$\Rightarrow \hat\beta_1 = \frac{\Sigma X_iY_i - n\bar X\bar Y}{\Sigma X_i^2 - n\bar X^2} = \frac{\Sigma(X_iY_i - \bar X\bar Y)}{\Sigma(X_i^2 - \bar X^2)}$$

$$= \frac{\Sigma(X_i - \bar X)(Y_i - \bar Y)}{\Sigma(X_i - \bar X)^2}$$

3.) f) To show $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ we can rewrite the right-hand side to match the form found by the matrix MLE:

$$\bar{Y} - \hat{\beta}_1 \bar{X} = \bar{Y} - \frac{\sum X_i Y_i - \bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2}\,\bar{X}$$

$$= \frac{\bar{Y}\left(\sum X_i^2 - n\bar{X}^2\right) - \bar{X}\left(\sum X_i Y_i - n\bar{X}\bar{Y}\right)}{\sum X_i^2 - n\bar{X}^2}$$

$$= \frac{\bar{Y}\sum X_i^2 - \cancel{n\bar{Y}\bar{X}^2} - \bar{X}\sum X_i Y_i + \cancel{n\bar{X}^2\bar{Y}}}{\sum X_i^2 - n\bar{X}^2}$$

$$= \frac{\bar{Y}\sum X_i^2 - \bar{X}\sum X_i Y_i}{\sum X_i^2 - n\bar{X}^2}$$