Solution to Practice Problems #4, #5 and #3q5.
For other problems, please refer to homework solutions and lectures

4. (a) I: $E[Y|X = x] = \beta_0 + \beta_1 x$.
II: $E[Y|Z = z] = \gamma_0 + \gamma_1 z = \gamma_0 + \gamma_1(ax + b) = (\gamma_0 + \gamma_1 b) + (\gamma_1 a)x$.
Thus we have $\beta_0 = \gamma_0 + \gamma_1 b$, $\beta_1 = \gamma_1 a$.
Comparison of Variances:
For I, since $\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$,

$$\hat{\sigma}_1^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \tag{1}$$

$$= \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \hat{\beta}_1(x_i - \overline{x}))^2. \tag{2}$$

For II, since $\hat{\gamma}_0 = \overline{y} - \hat{\gamma}_1 \overline{z}$,

$$\hat{\sigma}_2^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{\gamma}_0 - \hat{\gamma}_1 z_i)^2 \tag{3}$$

$$= \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \hat{\gamma}_1(z_i - \overline{z}))^2. \tag{4}$$

Since $z = ax + b$, we have $\overline{z} = a\overline{x} + b$, and then

$$z_i - \overline{z} = ax_i + b - (a\overline{x} + b) = a(x_i - \overline{x}). \tag{5}$$

Moreover,

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(z_i - \overline{z})}{\sum_{i=1}^{n}(z_i - \overline{z})^2} \tag{6}$$

$$= \frac{\sum_{i=1}^{n}(y_i - \overline{y})(ax_i + b - (a\overline{x} + b))}{\sum_{i=1}^{n}(ax_i + b - (a\overline{x} + b))^2} \tag{7}$$

$$= \frac{\sum_{i=1}^{n}(y_i - \overline{y}) \cdot a(x_i - \overline{x})}{\sum_{i=1}^{n}(a(x_i - \overline{x}))^2} \tag{8}$$

$$= \frac{1}{a}\frac{\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{9}$$

$$= \frac{1}{a}\hat{\beta}_1. \tag{10}$$

Plug (5) and (10) into (4), then we get

$$\hat{\sigma}_2^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \hat{\gamma}_1(z_i - \overline{z}))^2 \tag{11}$$
$$= \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \frac{1}{a}\hat{\beta}_1 \cdot a(x_i - \overline{x}))^2 \tag{12}$$
$$= \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \hat{\beta}_1(x_i - \overline{x}))^2 \tag{13}$$
$$= \hat{\sigma}_1^2. \tag{14}$$

Therefore, the estimates of variances are the same for I and II.

Comparison for t-tests:

For I, the test statistic for $\mathrm{H}_0 : \beta_1 = 0$ is

$$T_{\beta_1} = \frac{\hat{\beta}_1}{\hat{SE}(\hat{\beta}_1)},$$

where $\hat{SE}(\hat{\beta}_1)^2 = \frac{\hat{\sigma}_1^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$.

For II, the test statistic for $\mathrm{H}_0 : \gamma_1 = 0$ is

$$T_{\gamma_1} = \frac{\hat{\gamma}_1}{\hat{SE}(\hat{\gamma}_1)},$$

where $\hat{SE}(\hat{\gamma}_1)^2 = \frac{\hat{\sigma}_2^2}{\sum_{i=1}^{n}(z_i - \overline{z})^2} = \frac{\hat{\sigma}_1^2}{a^2\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{1}{a^2}\hat{SE}(\hat{\beta}_1)^2$.

Therefore, we have $\hat{SE}(\hat{\gamma}_1) = \frac{1}{a}\hat{SE}(\hat{\beta}_1)$.

By (10), $\hat{\gamma}_1 = \frac{1}{a}\hat{\beta}_1$, therefore we have

$$T_{\gamma_1} = \frac{\hat{\gamma}_1}{\hat{SE}(\hat{\gamma}_1)} = \frac{\frac{1}{a}\hat{\beta}_1}{\frac{1}{a}\hat{SE}(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{\hat{SE}(\hat{\beta}_1)} = T_{\beta_1},$$

the two t-test statistics are the same.

(b) I: $E[Y|X = x] = \beta_0 + \beta_1 x$.

III: $E[V|X = x] = \delta_0 + \delta_1 x$.

Since $E[V|X = x] = E[dY|X = x] = dE[Y|X = x] = d\beta_0 + d\beta_1 x$, we have $\delta_0 = d\beta_0$ and $\delta_1 = d\beta_1$.

Comparison of Variances:

Similarly as before, we have

$$\hat{\sigma}_3^2 = \frac{1}{n-2}\sum_{i=1}^{n}(v_i - \hat{v}_i)^2 = \frac{1}{n-2}\sum_{i=1}^{n}(v_i - \hat{\delta}_0 - \hat{\delta}_1 x_i)^2 \tag{15}$$
$$= \frac{1}{n-2}\sum_{i=1}^{n}(v_i - \overline{v} - \hat{\delta}_1(x_i - \overline{x}))^2. \tag{16}$$

Since $V = dY$, we have

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{n}(v_i - \overline{v})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(dy_i - d\overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = d\hat{\beta}_1.$$

Then

$$\begin{aligned}
\hat{\sigma}_3^2 &= \frac{1}{n-2}\sum_{i=1}^{n}(v_i - \overline{v} - \hat{\delta}_1(x_i - \overline{x}))^2 && (17)\\
&= \frac{1}{n-2}\sum_{i=1}^{n}(dy_i - d\overline{y} - d\hat{\beta}_1(x_i - \overline{x}))^2 && (18)\\
&= \frac{d^2}{n-2}\sum_{i=1}^{n}(y_i - \overline{y} - \hat{\beta}_1(x_i - \overline{x}))^2 && (19)\\
&= d^2\hat{\sigma}_1^2. && (20)
\end{aligned}$$

Comparison of t-tests:

For III, the test statistic for $\mathrm{H}_0 : \delta_1 = 0$ is

$$T_{\delta_1} = \frac{\hat{\delta}_1}{\hat{SE}(\hat{\delta}_1)},$$

where $\hat{SE}(\hat{\delta}_1)^2 = \frac{\hat{\sigma}_3^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = d^2\frac{\hat{\sigma}_1^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = d^2\hat{SE}(\hat{\beta}_1)^2$.

Therefore we have $\hat{SE}(\hat{\delta}_1) = d\hat{SE}(\hat{\beta}_1)$ and $T_{\delta_1} = \frac{d\hat{\beta}_1}{d\hat{SE}(\hat{\beta}_1)} = T_{\beta_1}$.

5. We consider $t = \frac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{S_{xx}}}$, where $S_{xx} = \sum_{i=1}^n (x_i - \overline{x})^2$. Similarly define $S_{yy} = \sum_{i=1}^n (y_i - \overline{y})^2$. For the sample correlation, we have

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^n (x_i - \overline{x})^2 \sum_{i=1}^n (y_i - \overline{y})^2}}.$$

Then

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2} = \sqrt{\frac{\sum_{i=1}^n (y_i - \overline{y})^2}{\sum_{i=1}^n (x_i - \overline{x})^2}} \cdot r_{xy} = \sqrt{\frac{S_{yy}}{S_{xx}}} \cdot r_{xy}.$$

Therefore we have

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \frac{1}{n-2} \sum_{i=1}^n (y_i - \overline{y} - \hat{\beta}_1 (x_i - \overline{x}))^2 \\
&= \frac{1}{n-2} \sum_{i=1}^n (y_i - \overline{y} - \sqrt{\frac{S_{yy}}{S_{xx}}} \cdot r_{xy} (x_i - \overline{x}))^2 \\
&= \frac{1}{n-2} [\sum_{i=1}^n (y_i - \overline{y})^2 - 2 \sum_{i=1}^n (y_i - \overline{y})(x_i - \overline{x}) \sqrt{\frac{S_{yy}}{S_{xx}}} \cdot r_{xy} + \sum_{i=1}^n (x_i - \overline{x})^2 \frac{S_{yy}}{S_{xx}} \cdot r_{xy}^2] \\
&= \frac{1}{n-2} [S_{yy} - 2\sqrt{S_{xx} \cdot S_{yy}} \cdot r_{xy} \sqrt{\frac{S_{yy}}{S_{xx}}} \cdot r_{xy} + S_{xx} \frac{S_{yy}}{S_{xx}} \cdot r_{xy}^2] \\
&= \frac{1}{n-2} [S_{yy} - 2S_{yy} \cdot r_{xy}^2 + S_{yy} \cdot r_{xy}^2] \\
&= \frac{1}{n-2} S_{yy} (1 - r_{xy}^2)
\end{aligned}$$

So $\hat{\sigma} = \frac{1}{\sqrt{n-2}} \sqrt{S_{yy}(1 - r_{xy}^2)}$.

Finally we have

$$t = \frac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{S_{xx}}} = \sqrt{n-2} \cdot \frac{1}{\sqrt{S_{yy}(1 - r_{xy}^2)}} \cdot \sqrt{\frac{S_{yy}}{S_{xx}}} \cdot r_{xy} \cdot \sqrt{S_{xx}} = \sqrt{n-2} \cdot \frac{r_{xy}}{\sqrt{1 - r_{xy}^2}}.$$

5. For a linear regression without an intercept, the ith fitted value is $\hat{y}_i = x_i\hat{\beta}$

where $\hat{\beta} = \dfrac{\sum_{i'=1}^{n} x_{i'}y_{i'}}{\sum_{i'=1}^{n} x_{i'}^2}$, we have $\hat{y}_i = x_i\hat{\beta}$

$$= \frac{\sum_{i'=1}^{n} x_{i'}y_{i'}}{\sum_{i'=1}^{n} x_{i'}^2 \rightarrow \text{a number}}\, x_i = \frac{\sum_{i'=1}^{n} x_{i'}x_i}{\sum_{i'=1}^{n} x_{i'}^2} \sum_{i'=1}^{n} y_{i'} = \text{(next line)}$$

$$= \sum_{i'=1}^{n} \left(\frac{x_{i'}x_i}{\sum_{i'=1}^{n} x_{i'}^2}\, y_{i'} \right) \text{①}$$

Let $a_{i'} = \dfrac{x_{i'}x_i}{\sum_{i'=1}^{n} x_{i'}^2}$, and ① becomes $\hat{y}_i = \sum_{i'=1}^{n} \left(\dfrac{x_{i'}x_i}{\sum_{i'=1}^{n} x_{i'}^2}\, y_{i'} \right) = \sum_{i'=1}^{n} a_{i'}y_{i'}$ □