1. Suppose $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ for some unknown $\sigma^2 > 0$. Let $\hat{\boldsymbol{y}} = \boldsymbol{P_X y}$. By the results on multivariate normal distributions from slide set 1 ("Preliminaries"), we have

$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}).$$

(a) Notice that we want to find the distribution of a linear transformation of $\boldsymbol{y}$:

$$\begin{bmatrix} \hat{\boldsymbol{y}} \\ \boldsymbol{y} - \hat{\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P_X y} \\ \boldsymbol{y} - \boldsymbol{P_X y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \boldsymbol{y}. \tag{1}$$

Together, slides 28 of the first slide set ("Preliminaries") imply that a linear transformation, say $\boldsymbol{Ax} + \boldsymbol{b}$, of a normal random variable $\boldsymbol{x}$ is also normal:

$$\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \boldsymbol{Ax} + \boldsymbol{b} \sim \mathcal{N}(\boldsymbol{A\mu} + \boldsymbol{b}, \boldsymbol{A\Sigma A}').$$

The mean of the linear transformation in (**??**), since $\boldsymbol{P_X X} = \boldsymbol{X}$, is

$$\begin{aligned}
\mathrm{E}\left(\begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \boldsymbol{y}\right) &= \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \mathrm{E}(\boldsymbol{y}) \\
&= \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \boldsymbol{X\beta} \\
&= \begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{X\beta} - \boldsymbol{X\beta} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix}.
\end{aligned}$$

The variance of (**??**), since $\boldsymbol{P_X}$ is symmetric and idempotent, is

$$\begin{aligned}
\mathrm{Var}\left(\begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \boldsymbol{y}\right) &= \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \mathrm{Var}(\boldsymbol{y}) \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix}' \\
&= \begin{bmatrix} \boldsymbol{P_X} \\ \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix} \sigma^2 \boldsymbol{I} \left[\boldsymbol{P_X}', \ (\boldsymbol{I} - \boldsymbol{P_X})'\right] \\
&= \sigma^2 \begin{bmatrix} \boldsymbol{P_X P_X}' & \boldsymbol{P_X}(\boldsymbol{I} - \boldsymbol{P_X})' \\ (\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{P_X}' & (\boldsymbol{I} - \boldsymbol{P_X})(\boldsymbol{I} - \boldsymbol{P_X})' \end{bmatrix} \\
&= \sigma^2 \begin{bmatrix} \boldsymbol{P_X} & \boldsymbol{P_X} - \boldsymbol{P_X} \\ \boldsymbol{P_X} - \boldsymbol{P_X} & \boldsymbol{I} - \boldsymbol{P_X} - \boldsymbol{P_X} + \boldsymbol{P_X} \end{bmatrix} \\
&= \sigma^2 \begin{bmatrix} \boldsymbol{P_X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix}.
\end{aligned}$$

As a linear transformation of a multivariate normal random variable, it follows that

$$\begin{bmatrix} \hat{\boldsymbol{y}} \\ \boldsymbol{y} - \hat{\boldsymbol{y}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{X\beta} \\ \boldsymbol{0} \end{bmatrix}, \ \sigma^2 \begin{bmatrix} \boldsymbol{P_X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} - \boldsymbol{P_X} \end{bmatrix}\right).$$

<u>Comments</u>: it is important to notice that the expected values of $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}$ depend on the parameter $\boldsymbol{\beta}$ and not the estimate $\hat{\boldsymbol{\beta}}$, which is a random variable. Similarly, the variances do not depend on the estimate $\hat{\sigma}^2$. Additionally, some students did not give the covariances or only stated that they were zero without any derivation (the distribution of a multivariate normal random variable is not fully specified without the covariances).

(b) Again using the fact that $\boldsymbol{P_X}$ is symmetric and idempotent, we see that $\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}}$ is a quadratic form:

$$\begin{aligned}
\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} &= [\boldsymbol{P_X}\boldsymbol{y}]'\boldsymbol{P_X}\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{P_X}'\boldsymbol{P_X}\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{P_X}\boldsymbol{P_X}\boldsymbol{y} \\
&= \boldsymbol{y}'\boldsymbol{P_X}\boldsymbol{y},
\end{aligned}$$

where $\boldsymbol{y} \sim \mathcal{N}\left(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I}\right)$. Recall the results about quadratic forms in the first slide set ("Preliminaries") on slide 31. To apply these, want to find a symmetric matrix $\boldsymbol{A}$ such that $\boldsymbol{A\Sigma}$ is idempotent for $\boldsymbol{\Sigma} \equiv \mathrm{Var}(\boldsymbol{y})$. We can't use $\boldsymbol{P_X}$ as our $\boldsymbol{A}$ matrix because the σ^2 doesn't cancel:

$$\boldsymbol{P_X}\,\mathrm{Var}(\boldsymbol{y}) = \boldsymbol{P_X}\sigma^2\boldsymbol{I} = \sigma^2\boldsymbol{P_X}.$$

Instead using $\boldsymbol{A} = \frac{\boldsymbol{P_X}}{\sigma^2}$, which is symmetric, gives

$$\boldsymbol{A\Sigma} = \boldsymbol{A}\,\mathrm{Var}(\boldsymbol{y}) = \frac{\boldsymbol{P_X}}{\sigma^2}\sigma^2\boldsymbol{I} = \boldsymbol{P_X},$$

which we know is idempotent.

It is easy to verify that $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$ is positive definite since $\sigma^2 > 0$. We need to determine the rank of $\boldsymbol{A}$:

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{P_X}/\sigma^2) = \mathrm{rank}(\boldsymbol{P_X}) = \mathrm{rank}(\boldsymbol{X}).$$

Then,

$$\frac{1}{\sigma^2}\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} = \boldsymbol{y}'\frac{\boldsymbol{P_X}}{\sigma^2}\boldsymbol{y} \sim \chi^2_{\mathrm{rank}(\boldsymbol{X})}\big([\boldsymbol{X\beta}]'\boldsymbol{AX\beta}/2\big),$$

where the noncentrality parameter simplifies to

$$\begin{aligned}
[\boldsymbol{X\beta}]'\boldsymbol{AX\beta}/2 &= [\boldsymbol{X\beta}]'\frac{\boldsymbol{P_X}}{\sigma^2}\boldsymbol{X\beta}\frac{1}{2} \\
&= \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{P_X}\boldsymbol{X\beta} \\
&= \frac{1}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta}.
\end{aligned}$$

Therefore, we end up with a scaled non-central chi-square random variable on $\mathrm{rank}(\boldsymbol{X})$ degrees of freedom:

$$\hat{\boldsymbol{y}}'\hat{\boldsymbol{y}} \sim \sigma^2\chi^2_{\mathrm{rank}(\boldsymbol{X})}\left(\frac{\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta}}{2\sigma^2}\right).$$

Comments: some students used results that apply to the sum of squared *independent* normal random variables with variance one (e.g., the result on slide 29 of set 1; this assumes $\boldsymbol{A} = \boldsymbol{\Sigma} = \boldsymbol{I}$). Others tried to find the distribution of the quadratic form $\hat{\boldsymbol{y}}\frac{\boldsymbol{I}}{\sigma^2}\hat{\boldsymbol{y}}$ rather than simplifying to $\boldsymbol{y}\frac{\boldsymbol{P_X}}{\sigma^2}\boldsymbol{y}$; unfortunately this doesn't work because $\text{Var}(\hat{\boldsymbol{y}}) = \sigma^2\boldsymbol{P_X}$ may not be positive definite.

2. (a) Note that

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \text{ and } \boldsymbol{W} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then, the column space of $\boldsymbol{X}$ is the same as the column space of $\boldsymbol{W}$ because there exist $\mathbf{C}_1$ and $\mathbf{C}_2$ such that

$$\boldsymbol{X} = \boldsymbol{W}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \boldsymbol{W}\mathbf{C}_1$$

and

$$\boldsymbol{W} = \boldsymbol{X}\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} = \boldsymbol{X}\mathbf{C}_2.$$

So, this model is the same as the cell means model. In addition to the more formal argument on the previous fact, it should be clear almost immediately that the model in question is the cell means model because $\beta_1, \beta_2, \beta_3$ and β_4 can be any real numbers. Also, $\beta_1, \beta_1+\beta_2, \beta_1+\beta_2+\beta_3$, and $\beta_1+\beta_2+\beta_3+\beta_4$ can be any real numbers with no restrictions. Thus, this model is the same as the cell means model.

(b) Let $\mu_1 = \beta_1, \mu_2 = \beta_1+\beta_2, \mu_3 = \beta_1+\beta_2+\beta_3, \mu_4 = \beta_1+\beta_2+\beta_3+\beta_4$. Then $\beta_4 = \mu_4 - \mu_3$ whose BLUE is $\bar{y}_{4.} - \bar{y}_{3.}$ under the cell means model. Thus, BLUE of β_4 is $\bar{y}_{4.} - \bar{y}_{3.} = 26.3 - 22.8 = 3.5$.

(c) Note that $Var\left(\hat{\beta}_4\right) = Var\left(\hat{\mu}_4 - \hat{\mu}_3\right) = Var\left(\bar{y}_{4.} - \bar{y}_{3.}\right) = \frac{\sigma^2}{4} + \frac{\sigma^2}{2}$ and

$$MSE = \frac{SSE}{11-4} = \frac{(3-1)4.1 + (2-1)3.4 + (2-1)2.8 + (4-1)3.2}{7} = \frac{8.2+3.4+2.8+9.6}{7} = \frac{24}{7}.$$

Thus, $SE(\hat{\beta}_4) = \sqrt{\hat{\sigma}^2\left(\frac{1}{4}+\frac{1}{2}\right)} = \sqrt{(MSE)\frac{3}{4}} = \sqrt{\frac{24}{7}\frac{3}{4}} = \sqrt{\frac{18}{7}}$.

3. (a) Simplify $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$.

$\boldsymbol{X} = \begin{bmatrix}\mathbf{1}, & \boldsymbol{x}\end{bmatrix}$ and $\boldsymbol{X}' = \begin{bmatrix}\mathbf{1}' \\ \boldsymbol{x}'\end{bmatrix}$,then

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= (\hat{\beta}_0,\, \hat{\beta}_1)' \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&= \left(\begin{bmatrix}\mathbf{1}' \\ \boldsymbol{x}'\end{bmatrix}\begin{bmatrix}\mathbf{1}, & \boldsymbol{x}\end{bmatrix}\right)^{-1}\begin{bmatrix}\mathbf{1}' \\ \boldsymbol{x}'\end{bmatrix}\boldsymbol{y} \\
&= \begin{bmatrix}\mathbf{1}'\mathbf{1} & \mathbf{1}'\boldsymbol{x} \\ \boldsymbol{x}'\mathbf{1} & \boldsymbol{x}'\boldsymbol{x}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{1}'\boldsymbol{y} \\ \boldsymbol{x}'\boldsymbol{y}\end{bmatrix} \\
&= \begin{bmatrix} n & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} x_i & \sum\limits_{i=1}^{n} x_i^2\end{bmatrix}^{-1}\begin{bmatrix}\sum\limits_{i=1}^{n} y_i \\ \sum\limits_{i=1}^{n} x_i y_i\end{bmatrix} \\
&= \frac{1}{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}\begin{bmatrix}\sum\limits_{i=1}^{n} x_i^2 & -\sum\limits_{i=1}^{n} x_i \\ -\sum\limits_{i=1}^{n} x_i & n\end{bmatrix}\begin{bmatrix}\sum\limits_{i=1}^{n} y_i \\ \sum\limits_{i=1}^{n} x_i y_i\end{bmatrix} \\
&= \frac{1}{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}\begin{bmatrix}\sum\limits_{i=1}^{n} x_i^2\sum\limits_{i=1}^{n} y_i - \sum\limits_{i=1}^{n} x_i\sum\limits_{i=1}^{n} x_i y_i \\ -\sum\limits_{i=1}^{n} x_i\sum\limits_{i=1}^{n} y_i + n\sum\limits_{i=1}^{n} x_i y_i\end{bmatrix}
\end{aligned}$$

Therefore,

$$\hat{\beta}_0 = \frac{\sum\limits_{i=1}^{n} x_i^2\sum\limits_{i=1}^{n} y_i - \sum\limits_{i=1}^{n} x_i\sum\limits_{i=1}^{n} x_i y_i}{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}$$

$$\hat{\beta}_1 = \frac{-\sum\limits_{i=1}^{n} x_i\sum\limits_{i=1}^{n} y_i + n\sum\limits_{i=1}^{n} x_i y_i}{n\sum\limits_{i=1}^{n} x_i^2 - (\sum\limits_{i=1}^{n} x_i)^2}$$

(b) Find a matrix $\boldsymbol{B}$ so that $\boldsymbol{X}\boldsymbol{B}^{-1} = \boldsymbol{W} = \begin{bmatrix}\mathbf{1}, & \boldsymbol{x} - \bar{x}.\mathbf{1}\end{bmatrix}$.

Notice that $\boldsymbol{B}$ is a 2×2 matrix,

$$\underset{n\times 2}{\boldsymbol{X}}\ \underset{2\times 2}{\boldsymbol{B}^{-1}} = \underset{n\times 2}{\boldsymbol{W}} \iff \boldsymbol{X} = \boldsymbol{W}\boldsymbol{B} \text{ for non-singular } \boldsymbol{B}$$

Take $\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$

$$\begin{aligned}
\boldsymbol{X} = \boldsymbol{W}\boldsymbol{B} &\iff \begin{bmatrix}\mathbf{1}, & \boldsymbol{x}\end{bmatrix} = \begin{bmatrix}\mathbf{1}, & \boldsymbol{x} - \bar{x}.\mathbf{1}\end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\
&\iff \begin{cases} \mathbf{1} = b_{11}\mathbf{1} + b_{21}(\boldsymbol{x} - \bar{x}.\mathbf{1}) \\ \boldsymbol{x} = b_{12}\mathbf{1} + b_{22}(\boldsymbol{x} - \bar{x}.\mathbf{1}) \end{cases} \\
&\iff \begin{cases} b_{11} - b_{21}\bar{x}. = 1 \\ b_{21} = 0 \\ b_{12} - b_{22}\bar{x}. = 0 \\ b_{22} = 1 \end{cases} \\
&\iff \begin{cases} b_{11} = 1 \\ b_{21} = 0 \\ b_{12} = \bar{x}. \\ b_{22} = 1 \end{cases}
\end{aligned}$$

Therefore $\boldsymbol{B} = \begin{bmatrix} 1 & \bar{x}. \\ 0 & 1 \end{bmatrix}$ and $\boldsymbol{B}^{-1} = \begin{bmatrix} 1 & -\bar{x}. \\ 0 & 1 \end{bmatrix}$.

(c) Derive expressions for the least squares estimators of α_0 and α_1 using $\hat{\boldsymbol{\alpha}} = (\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{y}$.

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} &= (\hat{\alpha_0},\, \hat{\alpha_1})' \\
&= (\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{y} \\
&= \left(\begin{bmatrix} \mathbf{1}' \\ \boldsymbol{x}' - \bar{x}.\mathbf{1}' \end{bmatrix} \begin{bmatrix} \mathbf{1}, & \boldsymbol{x} - \bar{x}.\mathbf{1} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{1}' \\ \boldsymbol{x}' - \bar{x}.\mathbf{1}' \end{bmatrix} \boldsymbol{y} \\
&= \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'(\boldsymbol{x} - \bar{x}.\mathbf{1}) \\ (\boldsymbol{x}' - \bar{x}.\mathbf{1}')\mathbf{1} & (\boldsymbol{x}' - \bar{x}.\mathbf{1}')(\boldsymbol{x} - \bar{x}.\mathbf{1}) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'\boldsymbol{y} \\ \boldsymbol{x}'\boldsymbol{y} - \bar{x}.\mathbf{1}'\boldsymbol{y} \end{bmatrix} \\
&= \begin{bmatrix} n & \sum_{i=1}^{n} x_i - n\bar{x}. \\ \sum_{i=1}^{n} x_i - n\bar{x}. & \sum_{i=1}^{n} x_i^2 - n\bar{x}.^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i - \bar{x}. \sum_{i=1}^{n} y_i \end{bmatrix} \\
&= \begin{bmatrix} n & 0 \\ 0 & \sum_{i=1}^{n} x_i^2 - n\bar{x}.^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} (x_i - \bar{x}.) y_i \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} (x_i - \bar{x}.) y_i \end{bmatrix} \\
&= \begin{bmatrix} \frac{\sum_{i=1}^{n} y_i}{n} \\ \frac{\sum_{i=1}^{n} (x_i - \bar{x}.) y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2} \end{bmatrix}
\end{aligned}$$

(d) Multiply $\hat{\boldsymbol{\alpha}}$ from part (c) by $\boldsymbol{B}^{-1}$ from part (b) to obtain expressions for $\hat{\beta}_0$ and $\hat{\beta}_1$.

$$\begin{aligned}
(\hat{\beta}_0,\, \hat{\beta}_1)' &= \boldsymbol{B}^{-1}\hat{\boldsymbol{\alpha}} \\
&= \begin{bmatrix} 1 & -\bar{x}. \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sum_{i=1}^{n} y_i}{n} \\ \frac{\sum_{i=1}^{n} (x_i - \bar{x}.) y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2} \end{bmatrix} \\
&= \begin{bmatrix} \bar{y}. - \bar{x}. \frac{\sum_{i=1}^{n} (x_i - \bar{x}.) y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2} \\ \frac{\sum_{i=1}^{n} (x_i - \bar{x}.) y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2} \end{bmatrix}
\end{aligned}$$

(e) Show that your answer to part (a) matches your answer to part (d).

Notice that $\sum_{i=1}^{n} y_i = n\bar{y}.$ and $\sum_{i=1}^{n} x_i = n\bar{x}..$

$$\text{In part (d), } \hat{\beta}_0 = \bar{y}. - \bar{x}. \frac{\sum_{i=1}^{n}(x_i - \bar{x}.)y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2}$$

$$= \frac{\bar{y}.(\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2) - \bar{x}.(\sum_{i=1}^{n} x_iy_i - \bar{x}. \sum_{i=1}^{n} y_i)}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2}$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i \cancel{- \bar{x}.^2 \sum_{i=1}^{n} y_i} - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_iy_i \cancel{+ \bar{x}.^2 \sum_{i=1}^{n} y_i}}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_iy_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \hat{\beta}_0 \text{ in part (a).}$$

$$\text{In part (d), } \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}.)y_i}{\sum_{i=1}^{n} x_i^2 - n\bar{x}.^2}$$

$$= \frac{\sum_{i=1}^{n} x_iy_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2}$$

$$= \frac{-\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + n\sum_{i=1}^{n} x_iy_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \hat{\beta}_1 \text{ in part (a).}$$

4. (a) Let Jane be instructor 1 and John be instructor 2. For $i = 1, 2$ and $j = 1, 2$, let μ_{ij} be the expected value of a rating given by a student whose actual instructor was i and whose perceived instructor was j. The best estimates of $\mu_{11}, \mu_{12}, \mu_{21}$, and μ_{22} are provided by R as

$$\begin{aligned}
\hat{\mu}_{11} &= 2.85 \\
\hat{\mu}_{12} &= 2.85 + 0.87 \\
\hat{\mu}_{21} &= 2.85 - 0.06 \\
\hat{\mu}_{22} &= 2.85 - 0.06 + 0.87 - 0.1831
\end{aligned}$$

Thus, 0.87 is an estimator of $\mu_{12} - \mu_{11}$. This means that when Jane is the actual instructor, the expected value of her rating is estimated to be 0.87 points higher when students believe she is John rather than when students are told the truth about her identity. The standard error of 0.3184 says that the "typical size" of the error made when estimating $\mu_{12} - \mu_{11}$ (using the method we've used here) is approximately 0.3184 units. The p-value of 0.00941 indicates that an estimated difference as large or larger than 0.87 would be unlikely to occur if μ_{12} were equal to μ_{11}. Thus, Jane was rated significantly higher by the students who believed she was John than by students who knew she was Jane.

(b) The relevant estimate is

$$\begin{aligned} \frac{\hat{\mu}_{11}+\hat{\mu}_{21}}{2} - \frac{\hat{\mu}_{12}+\hat{\mu}_{22}}{2} &= \frac{2.85+2.85-0.06}{2} - \frac{2.85+0.87+2.85-0.06+0.87-0.1831}{2} \\ &= -0.87 + \frac{0.1831}{2}. \end{aligned}$$

The variance of the estimator is

$$\begin{aligned} \frac{1}{4}\left[Var\left(\hat{\mu}_{11}\right) + Var\left(\hat{\mu}_{21}\right) + Var\left(\hat{\mu}_{12}\right) + Var\left(\hat{\mu}_{22}\right)\right] &= \frac{1}{4}\left[\frac{\sigma^2}{10} + \frac{\sigma^2}{10} + \frac{\sigma^2}{10} + \frac{\sigma^2}{13}\right] \\ &= \frac{\sigma^2}{4}\left(\frac{3}{10} + \frac{1}{13}\right). \end{aligned}$$

We will need $\hat{\sigma}^2$. Note $0.2252 = SE\left(\hat{\mu}_{11}\right) = \sqrt{\hat{\sigma}^2/10}$. Thus, $\hat{\sigma}^2 = 10(0.2252)^2$. The confidence interval is

$$-0.87 + \frac{0.1831}{2} \pm t_{43-4,0.975}\sqrt{\frac{10(0.2252)^2}{4}\left(\frac{3}{10} + \frac{1}{13}\right)}.$$