Jared Yu
Test 1

1. Assume that a linear regression model has three regressors.
   a. Derive a test statistic to test the null hypothesis that their respective regression coefficients are all zero.

Ans:
A linear regression model with $k = 3$ regressors appears as follows,
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon.$$
A hypothesis test that is checking if the regression coefficients are all zero would appear as follows,
$$H_0: \beta_1 = \beta_2 = \beta_3 = 0 \text{ vs.} H_1: \beta_j \neq 0 \text{ for at least one } j \text{ for } j \in \{1,2,3\}.$$
The next step is to derive a test statistic for the above hypothesis. The test statistic used is the F_0 test statistic which appears as follows,
$$F_0 = \frac{\frac{SS_R}{k}}{\frac{SS_{Res}}{n-k-1}} = \frac{MS_R}{MS_{Res}},$$
where
$$MS_R = \frac{SS_R}{k} = \frac{1}{k}\left[\widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \frac{(\sum_{i=1}^n y_i)^2}{n}\right],$$
$$MS_{Res} = \frac{SS_{Res}}{n-k-1} = \frac{1}{n-k-1}[\mathbf{y}'\mathbf{y} - \widehat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}],$$
$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$
In the above, it is also the case that there is a total of n observations, for y_i, $i = 1, \cdots, n$. Another assumption is that $(\mathbf{X}'\mathbf{X})^{-1}$ exists, in other words $\mathbf{X}'\mathbf{X}$ is invertible. This requires that the columns of $\mathbf{X}$ are linearly independent. Here, $\mathbf{X}$ is the design matrix that includes the variables x_1, x_2, and x_3, in addition to a column of 1's to match with the intercept term β_0.

The textbook also states that
$$E(MS_{Res}) = \sigma^2 \text{ and } E(MS_R) = \sigma^2 + \frac{\boldsymbol{\beta}^{*'}\mathbf{X}_c'\mathbf{X}_c\boldsymbol{\beta}^*}{k\sigma^2},$$
where $\boldsymbol{\beta}^{*'} = (\beta_1, \beta_2, \beta_3)'$ and $\mathbf{X}_c$ is the "centered" model matrix. The $\mathbf{X}_c$ can be seen as follows,
$$\mathbf{X}_c = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & x_{13} - \bar{x}_3 \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & x_{23} - \bar{x}_3 \\ \vdots & \vdots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & x_{n3} - \bar{x}_3 \end{bmatrix}.$$
It can be seen that each of the columns have their values subtracted by their respective sample means.

If the test statistic F_0 is large, then that indicates that $\frac{\boldsymbol{\beta}^{*'}\mathbf{X}_c'\mathbf{X}_c\boldsymbol{\beta}^*}{k\sigma^2}$ is large. This implies that it is likely that at least one of $\beta_j \neq 0$, for $j \in \{1,2,3\}$. The test statistic F_0 actually follows a noncentral F distribution with degrees of freedom k and $n-k-1$ that has a noncentrality parameter
$$\lambda = \frac{\boldsymbol{\beta}^{*'}\mathbf{X}_c'\mathbf{X}_c\boldsymbol{\beta}^*}{k\sigma^2}.$$

The noncentrality parameter λ indicates that the test statistic F_0 should be large if at least one $\beta_j \neq 0$ for $j \in \{1,2,3\}$. To test the hypothesis, then the null can be rejected if $F_0 > F_{\alpha,k,n-k-1}$, where α is the confidence level for the hypothesis test.

b. If the null hypothesis in a) is rejected, would this mean that the regression model has value for prediction?

Ans:

The rejection of the null hypothesis stated above implies that at least one of the regressors, β_1, β_2, and β_3 contribute significantly to the model. This implies that for example out of x_1, x_2, and x_3 at least one of them is a useful regression coefficient at predicting y. This does not necessarily mean that the current regression model is good. For example, it is possible that only β_1 is nonzero and the others are roughly zero. In that case, it'd possibly make sense to slim down the model to a simple linear regression model involving only x_1 as the regressor. In such a case, the regression model with $k = 3$ regressors could possibly provide some value, but it would definitely benefit from some refinement.

c. Can the importance of each regressor be assessed by simply looking at the respective magnitudes of the t-statistics?

Ans:

If the hypothesis test in part b) is truly used and the result is such that the null is rejected, then the next step is to try and figure out which of the coefficients are worth keeping in the model. In such a problem, it'd help to create individual hypothesis tests as follows,

$$H_0: \beta_j = 0 \text{ vs. } H_1: \beta_j \neq 0 \text{ for } j = 1, 2, 3.$$

This would test each of the variables x_1, x_2, and x_3 to see if they can be removed from the model. If we fail to reject the null hypothesis for any of the j, $j \in \{1,2,3\}$, then that variable can be deleted from the model.

The test statistic is as follows,

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} = \frac{\hat{\beta}_j}{\mathrm{se}(\hat{\beta}_j)},$$

where C_{jj} is the diagonal element of $(\mathbf{X}'\mathbf{X})^{-1}$ that corresponds to $\hat{\beta}_j$ (*Note*: it is assumed that $(\mathbf{X}'\mathbf{X})^{-1}$ is calculatable, since it was used previously to calculate the β_j terms). The null hypothesis is rejected in the case if $|t_0| > t_{\frac{\alpha}{2},n-k-1}$ at confidence level α (*Note:* Here, t is the Student's t-distribution with degree of freedom $n-k-1$. Also, here $k = 3$.). Furthermore, this test is looking at the contribution of the variable x_j, when the other variables are included in the model. Therefore, it is not so straightforward to say that a regressor is good simply because it has a large test statistic t_0. It must also be understood that it is within the context of all the other regressors included.

State the assumptions for each step of your discussion or derivation in a), b), c).

2. In a simple linear regression analysis where the regressor x is non-random, we have random errors and residuals.
    a. What are the differences between residuals and random errors?

Ans:
The simple linear regression is as follows,

$$y = \beta_0 + \beta_1 x + \varepsilon,$$

where the intercept β_0 and the slope β_1 are unknown constants and ε is a random error component. The errors are assumed to have mean zero and unknown variance σ^2. Additionally, an assumption is that the errors are uncorrelated. This term, ε, helps to complete the linear relationship between the x and y observations in the data. In general, there will not be a perfect linear fit, and so the random error helps to compensate for the difference between y and $\beta_0 + \beta_1 x$.

Then, for n pairs of data $(y_1, x_1), \cdots, (y_n, x_n)$ the sample regression model is,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i.$$

Applying the method of least squares, the intercept, β_0, and slope, β_1, can be estimated with

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \text{ and } \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. The fitted simple linear regression model can be written as,

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x.$$

After having calculated the fitted model, the residual is calculated as follows,

$$e_i = y_i - \hat{y}_i = y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right), \qquad \text{for } i = 1, \cdots, n.$$

It is the difference between the observed y values and the fitted value for that sample. These residuals are helpful in the checking the model adequacy. By observing the behavior of the residuals, it can be a way to see if there is a departure from the underlying assumptions for the model.

Compared to the random error, the residual is like the realization and the random error is the random variable. For example, if there is a random variable X that represents the result of a die roll, then $x = 6$ could be the realization of that random variable after a single roll of the die. So, from the theoretical model, there is ε, but in the application of the model in a sample, there is e. It can be seen then that ε is like an unknown population variable and e is like an observable variable based on a sample of data.

b. Derive the variance of residual and the variance of random error.

Ans:
The variance of the residual is as follows:

$$Var(e_i) = Var(y_i - \hat{y}_i) = Var\left(y_i - \left(\hat{\beta}_0 + \hat{\beta}_1 x_i\right)\right)$$
$$= Var\left(y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i\right) = Var\left(y_i - \bar{y} + \hat{\beta}_1(\bar{x} - x_i)\right)$$
$$= Var\left(y_i - \bar{y} + \left(\frac{\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})}{\sum_{k=1}^{n}(x_k - \bar{x})^2}\right)(\bar{x} - x_i)\right)$$
$$= Var\left(y_i - \bar{y} + \left(\sum_{j=1}^{n} c_j y_j\right)(\bar{x} - x_i)\right)$$

Note: $c_i = \frac{x_i - \bar{x}}{\sum_{j=1}^{n}(x_j-\bar{x})^2}$, the reason is that the $x's$ are fixed and so they are considered constants and so the c_i term can come out of $\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{j=1}^{n}(x_j-\bar{x})^2}$. The denominator can be treated as a constant term. The numerator can be rewritten as follows,

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

Going back to the previous equation…

$$Var\left(y_i - \bar{y} + \left(\sum_{j=1}^{n} c_j y_j\right)(\bar{x}-x_i)\right)$$
$$= Var(y_i-\bar{y}) + Var\left[\sum_{j=1}^{n} c_j y_j\,(\bar{x}-x_i)\right] + 2Cov\left(y_i-\bar{y}, \sum_{j=1}^{n} c_j y_j\,(\bar{x}-x_i)\right)$$
$$= Var(y_i) + Var(\bar{y}) - 2Cov(y_i,\bar{y}) + (\bar{x}-x_i)^2 Var\left(\sum_{j=1}^{n} c_j y_j\right)$$
$$+ 2Cov\left(y_i, \sum_{j=1}^{n} c_j y_j\,(\bar{x}-x_i)\right) - 2Cov\left(\bar{y}, \sum_{j=1}^{n} c_j y_j\,(\bar{x}-x_i)\right)$$
$$= \sigma^2 + \frac{\sigma^2}{n} - 2Cov\left(y_i, \frac{y_i}{n}\right) + (\bar{x}-x_i)^2 \sum_{j=1}^{n} c_j^2 Var(y_j) + 2(\bar{x}-x_i)Cov(y_i, c_i y_i)$$
$$- 2(\bar{x}-x_i)\sum_{j=1}^{n} Cov\left(\frac{y_i}{n}, c_j y_j\right)$$
$$= \sigma^2 + \frac{\sigma^2}{n} - \frac{2}{n}\sigma^2 + (\bar{x}-x_i)^2 \frac{\sigma^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2} + 2(\bar{x}-x_i)c_i\sigma^2 - 2(\bar{x}-x_i)\sum_{j=1}^{n}\frac{c_j}{n}\sigma^2$$
$$= \sigma^2 - \frac{\sigma^2}{n} - \frac{(\bar{x}-x_i)^2\sigma^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2} + \frac{2(\bar{x}-x_i)^2\sigma^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2} - 0$$
$$= \sigma^2\left(1-\frac{1}{n}\right) + \frac{(\bar{x}-x_i)^2\sigma^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2}$$
$$\boxed{= \sigma^2\left[1 - \frac{1}{n} + \frac{(\bar{x}-x_i)^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2}\right]}$$

Note: An important fact used above is that the y_i terms are uncorrelated and so the covariance between y_i and y_j for $i \neq j$ is 0.

The above equation shows that the variance of the residuals can be written as follows,

$$Var(e_i) = \sigma^2\left[1 - \frac{1}{n} + \frac{(\bar{x} - x_i)^2}{\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}\right].$$

In the above derivation, some assumptions used are that the y_i terms are independent which is important within the variance and covariance functions. Also, the x terms are treated as regular constants and not as random variables.

In simple linear regression, some assumptions are that β_0 and β_1 are unknown constants, while the random error ε is assumed to have mean zero and unknown variance σ^2. The random error is also often said to follow a normal distribution, $N(0, \sigma^2)$. This is an idealized version of the simple linear regression model and it is not always the case that this property holds within a sample of data. So, it makes sense to say that the variance of the random error is σ^2, but it can't be shown in a mathematical manner.

c. Now assume x is random. Derive the variance of residual.

Ans:

In the case that x is random, there are different steps required in order to find out the variance of the residual. Starting from $Var(e_i) = Var(y_i - \hat{y}_i)$, it is possible to calculate up to the following step,

$$Var(y_i - \bar{y}) + Var\left[\sum_{j=1}^{n} c_j y_j\,(\bar{x} - x_i)\right] + 2Cov\left(y_i - \bar{y}, \sum_{j=1}^{n} c_j y_j\,(\bar{x} - x_i)\right).$$

The reason it stops here is that since the x terms are random, they can't be easily separated within the c_j term like they were before. So, the derivation for the variance of the residual will begin from another point. It will try the variance based on the formula for the law of total variance.

$$Var(e_i) = E[Var(e_i|x)] + Var[E(e_i|x)]$$

Note: Here, x is a random variable. Let us first look at $E(e_i|x)$.

$$E(e_i|x) = E(y_i - \bar{y}|x) = E(y_i|x) - E(\bar{y}|x)$$
$$= \beta_0 + \beta_1 x_i - \beta_0 - \beta_1\bar{x} = \beta_1(x_i - \bar{x})$$

Now, let us look at $Var[E(e_i|x)]$.

$$Var[E(e_i|x)] = Var[\beta_1(x_i - \bar{x})] = \beta_1^2[Var(x_i) + Var(\bar{x}) - 2Cov(x_i, \bar{x})]$$
$$= \beta_1^2\left[\sigma_x^2 + \frac{\sigma_x^2}{n} - \frac{2\sigma_x^2}{n}\right] = \beta_1^2\sigma_x^2\left(1 - \frac{1}{n}\right)$$

Note: An assumption is that x and y are jointly distributed, but the joint distribution is unknown. Also, the $x's$ are independent random variables whose probability distribution doesn't include β_0, β_1, or σ^2. Also, let σ_x^2 represent the variance of the random variable x.

Next, let us look at $Var(e_i|x)$. From part b), it has been shown that this evaluates to,

$$Var(e_i|x) = \sigma^2\left[1 - \frac{1}{n} + \frac{(\bar{x} - x_i)^2}{\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}\right].$$

Then, taking the expectation we have,

$$E[Var(e_i|x)] = \sigma^2\left\{1 - \frac{1}{n} + E\left[\frac{(\bar{x} - x_i)^2}{\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}\right]\right\}.$$

The expectation can't be simplified further, since there is not enough information about x, for example what sort of distribution it follows.

Finally, going back to $Var(e_i)$ and combining the terms,

$$Var(e_i) = E[Var(e_i|x)] + Var[E(e_i|x)]$$
$$= \boxed{\sigma^2 \left\{1 - \frac{1}{n} + E\left[\frac{(\bar{x} - x_i)^2}{\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}\right]\right\} + \beta_1^2 \sigma_x^2 \left(1 - \frac{1}{n}\right)}.$$

This formula is quite similar to the previous version in part b), however it includes some additional complexity due to x being a random variable.

State the assumptions for each step of your discussion or derivation in a), b), c).

3. In a multiple linear regression analysis, the response variable y is studied with two non-random variables x_1 and x_2. This regression model that is fitted to the data on (y, x_1, x_2) of n subjects is given by
$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \varepsilon,$$
where $z_1 = w_1 x_1 + (1 - w_1)x_2$, $z_2 = w_2 x_1 + (1 - w_2)x_2$, ε is a random error with mean zero and variance σ^2 (its value is unknown), and the weights w_1 and w_2 have known values.
   a. Derive the ordinary least-squares estimators for $(\beta_0, \beta_1, \beta_2)'$ as functions of (y, x_1, x_2), not (y, z_1, z_2).

Ans: (*Note: Some of the answer is borrowed from a discussion question that utilizes the same methodology to derive the OLS estimators in a multiple linear regression model with $k = 2$ regressors.*)
Let $\mathbf{Z}$ be the design matrix for $\boldsymbol{Z}$, the data matrix with dimensions $(n \times (k = 2))$. It has dimensions $n \times p$, where $p = (k = 2) + 1 = 3$, since it includes the column of 1's in the first position. Let $\mathbf{y}$ be the $n \times 1$ vector of the observations. An assumption for OLS is that $n < k$, $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$, and that the errors are uncorrelated.

Using the least-squares normal equations, the OLS estimates of β_i for $i = 0,1,2$, are as follows,

$$\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}.$$

This formula is based on the least-squares normal equations, $\mathbf{Z}'\boldsymbol{Z}\widehat{\boldsymbol{\beta}} = \mathbf{Z}'\boldsymbol{y}$, which can be found by minimizing the following,

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{Z}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{Z}\boldsymbol{\beta}).$$

An assumption for this is that $(\mathbf{Z}'\mathbf{Z})^{-1}$ exists, which is possible if the regressors (i.e., z_1 and z_2) are linearly independent. The next step is to find $(\mathbf{Z}'\mathbf{Z})^{-1}$. Let

$$\mathbf{Z} = \begin{bmatrix} 1 & z_{11} & z_{12} \\ \vdots & \vdots & \vdots \\ 1 & z_{n1} & z_{n2} \end{bmatrix},$$

then

$$\mathbf{Z'Z} = \begin{bmatrix} n & \sum_{i=1}^{n} z_{i1} & \sum_{i=1}^{n} z_{i2} \\ \sum_{i=1}^{n} z_{i1} & \sum_{i=1}^{n} z_{i1}^2 & \sum_{i=1}^{n} z_{i1}z_{i2} \\ \sum_{i=1}^{n} z_{i2} & \sum_{i=1}^{n} z_{i2}z_{i1} & \sum_{i=1}^{n} z_{i2}^2 \end{bmatrix}$$

Note: To simplify notation, allow $\sum_{i=1}^{n}(\cdot)$ to be shortened to $\sum(\cdot)$.

Furthermore, let $SSZ_1 = \sum z_{i1}^2 - \frac{(\sum z_{i1})^2}{n}, SSZ_2 = \sum z_{i2}^2 - \frac{(\sum z_{i2})^2}{n}, SSZ_{12} = \sum z_{i1}z_{i2} - \frac{\sum z_{i1}\sum z_{i2}}{n}$. The following steps will attempt to find the inverse of $\mathbf{Z'Z}$. The row operations will be abbreviated with R1, R2, and R3.

$$\left\langle \begin{array}{ccc|ccc} n & \sum z_{i1} & \sum z_{i2} & 1 & 0 & 0 \\ \sum z_{i1} & \sum z_{i1}^2 & \sum z_{i1}z_{i2} & 0 & 1 & 0 \\ \sum z_{i2} & \sum z_{i2}z_{i1} & \sum z_{i2}^2 & 0 & 0 & 1 \end{array} \right\rangle$$

Divide R1 by n:

$$\left\langle \begin{array}{ccc|ccc} 1 & \frac{\sum z_{i1}}{n} & \frac{\sum z_{i2}}{n} & \frac{1}{n} & 0 & 0 \\ \sum z_{i1} & \sum z_{i1}^2 & \sum z_{i1}z_{i2} & 0 & 1 & 0 \\ \sum z_{i2} & \sum z_{i2}z_{i1} & \sum z_{i2}^2 & 0 & 0 & 1 \end{array} \right\rangle$$

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ \sum z_{i1} & \sum z_{i1}^2 & \sum z_{i1}z_{i2} & 0 & 1 & 0 \\ \sum z_{i2} & \sum z_{i2}z_{i1} & \sum z_{i2}^2 & 0 & 0 & 1 \end{array} \right\rangle$$

R2 – R1*$\sum z_{i1}$:

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & \sum z_{i1}^2 - \frac{(\sum z_{i1})^2}{n} & \sum z_{i1}z_{i2} - \frac{\sum z_{i1}\sum z_{i2}}{n} & -\bar{z}_1 & 1 & 0 \\ \sum z_{i2} & \sum x_{i2}x_{i1} & \sum z_{i2}^2 & 0 & 0 & 1 \end{array} \right\rangle$$

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & SSZ_1 & SSZ_{12} & -\bar{z}_1 & 1 & 0 \\ \sum z_{i2} & \sum z_{i2}z_{i1} & \sum z_{i2}^2 & 0 & 0 & 1 \end{array} \right\rangle$$

R3 – R1*$\sum z_{i2}$:

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{x}_2 & \frac{1}{n} & 0 & 0 \\ 0 & SSZ_1 & SSZ_{12} & -\bar{z}_1 & 1 & 0 \\ 0 & \sum z_{i2}z_{i1} - \frac{\sum z_{i1}\sum z_{i2}}{n} & \sum z_{i2}^2 - \frac{(\sum z_{i2})^2}{n} & -\frac{\sum z_{i2}}{n} & 0 & 1 \end{array} \right\rangle$$

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & SSZ_1 & SSZ_{12} & -\bar{z}_1 & 1 & 0 \\ 0 & SSZ_{12} & SSZ_2 & -\bar{z}_2 & 0 & 1 \end{array}\right\rangle$$

Divide R2 by SSZ_1:

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & \frac{SSZ_{12}}{SSZ_1} & \frac{-\bar{z}_1}{SSZ_1} & \frac{1}{SSZ_1} & 0 \\ 0 & SSZ_{12} & SSZ_2 & -\bar{z}_2 & 0 & 1 \end{array}\right\rangle$$

R3 – R2 * SSZ_{12}:

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & \frac{SSZ_{12}}{SSZ_1} & \frac{-\bar{z}_1}{SSZ_1} & \frac{1}{SSZ_1} & 0 \\ 0 & 0 & SSZ_2-\frac{SSZ_{12}^2}{SSZ_1} & -\bar{z}_2+\frac{\bar{z}_1 SSZ_{12}}{SSZ_1} & -\frac{SSZ_{12}}{SSZ_1} & 1 \end{array}\right\rangle$$

Divide R3 by $SSZ_2-\frac{SSZ_{12}^2}{SSZ_1}$:

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & \frac{SSZ_{12}}{SSZ_1} & \frac{-\bar{z}_1}{SSZ_1} & \frac{1}{SSZ_1} & 0 \\ 0 & 0 & 1 & c & \frac{-SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} & \frac{SSZ_1}{SSZ_1\times SSZ_2-SSZ_{12}^2} \end{array}\right\rangle$$

$$c=\frac{-\bar{z}_2+\frac{\bar{z}_1 SSZ_{12}}{SSZ_1}}{SSZ_2-\frac{SSZ_{12}^2}{SSZ_1}}=\frac{-\bar{z}_2 SSZ_1+\bar{z}_1 SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2}$$

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & \frac{SSZ_{12}}{SSZ_1} & \frac{-\bar{z}_1}{SSZ_1} & \frac{1}{SSZ_1} & 0 \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1+\bar{z}_1 SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} & \frac{SSZ_1}{SSZ_1\times SSZ_2-SSZ_{12}^2} \end{array}\right\rangle$$

R2 – R3 * $\frac{SSZ_{12}}{SSZ_1}$:

$$\left\langle\begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & 0 & e & f & -\frac{SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1+\bar{z}_1 SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2} & \frac{SSZ_1}{SSZ_1\times SSZ_2-SSZ_{12}^2} \end{array}\right\rangle$$

$$\mathrm{e}=\frac{-\bar{z}_1}{SSZ_1}-\frac{-\bar{z}_2 SSZ_1+\bar{z}_1 SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2}\times\frac{SSZ_{12}}{SSZ_1}\frac{-1}{SSZ_1}=\left[\bar{z}_1+\frac{-\bar{z}_2 SSZ_1 SSZ_{12}+\bar{z}_1 SSZ_{12}^2}{SSZ_1\times SSZ_2-SSZ_{12}^2}\right]$$

$$=\frac{-1}{SSZ_1}\left[\frac{\bar{z}_1 SSZ_1 SSZ_2-\bar{z}_2 SSZ_1 SSZ_{12}}{SSZ_1\times SSZ_2-SSZ_{12}^2}\right]=\frac{\bar{z}_2 SSZ_{12}-\bar{z}_1 SSZ_2}{SSZ_1\times SSZ_2-SSZ_{12}^2}$$

$$\mathrm{f} = \frac{1}{SSZ_1} - \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \times \frac{SSZ_{12}}{SSZ_1} = \frac{1}{SSZ_1}\left[1 + \frac{SSZ_{12}^2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\right]$$
$$= \frac{1}{SSZ_1}\left[\frac{SSZ_1 \times SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\right] = \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}$$

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & \bar{z}_2 & \frac{1}{n} & 0 & 0 \\ 0 & 1 & 0 & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \end{array} \right\rangle$$

R1 – R3*$\bar{z}_2$:

$$\left\langle \begin{array}{ccc|ccc} 1 & \bar{z}_1 & 0 & \frac{1}{n} - \frac{-\bar{z}_2^2 SSZ_1 + \bar{z}_1 \bar{z}_2 SSZ_{12}}{SSX_1 \times SSX_2 - SSX_{12}^2} & \frac{\bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{\bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 1 & 0 & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \end{array} \right\rangle$$

R1 – R2*$\bar{z}_1$:

$$\left\langle \begin{array}{ccc|ccc} 1 & 0 & 0 & \mathrm{g} & \mathrm{h} & \mathrm{i} \\ 0 & 1 & 0 & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \end{array} \right\rangle$$

$$\mathrm{g} = \frac{1}{n} - \frac{-\bar{z}_2^2 SSZ_1 + \bar{z}_1 \bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} - \frac{\bar{z}_1 \bar{z}_2 SSZ_{12} - \bar{z}_1^2 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}$$
$$= \frac{1}{n} - \frac{-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1 \bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}$$
$$\mathrm{h} = \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}$$
$$\mathrm{i} = \frac{\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}$$

$$\left\langle \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{n} - \frac{-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1 \bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 1 & 0 & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ 0 & 0 & 1 & \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \end{array} \right\rangle$$

The following has been shown:

$$(\mathbf{Z'Z})^{-1} = \begin{bmatrix} \frac{1}{n} - \frac{-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1\bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \\ \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & \frac{-SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} & -\frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \end{bmatrix}$$

Where

$$SSZ_1 = \sum z_{i1}^2 - \frac{(\sum z_{i1})^2}{n}, SSZ_2 = \sum z_{i2}^2 - \frac{(\sum z_{i2})^2}{n}, SSZ_{12} = \sum z_{i1} z_{i2} - \frac{\sum z_{i1} \sum z_{i2}}{n}, \bar{z}_1 = \frac{\sum z_{i1}}{n}, \bar{z}_2 = \frac{\sum z_{i2}}{n}.$$

Then, from equation $\hat{\boldsymbol{\beta}} = (\mathbf{Z'Z})^{-1}\mathbf{Z'y}$:

$$\mathbf{Z'y} = \begin{bmatrix} 1 & \cdots & 1 \\ z_{11} & \cdots & z_{n1} \\ z_{12} & \cdots & z_{n2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum z_{i1} y_i \\ \sum z_{i2} y_i \end{bmatrix}$$

From this result it follows that for $\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = (\mathbf{Z'Z})^{-1}\mathbf{Z'y}$:

$$\hat{\beta}_0 = \frac{\sum y_i}{n} - \frac{-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1\bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum y_i + \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i1} y_i + \frac{\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i2} y_i$$

$$\hat{\beta}_1 = \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum y_i + \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i1} y_i - \frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i2} y_i$$

$$\hat{\beta}_2 = \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum y_i - \frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i1} y_i + \frac{SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2} \sum z_{i2} y_i$$

Before rewriting the above terms in the form of x_1 and x_2, the separate variables will be changed first.

$$\bar{z}_1 = w_1 \bar{x}_1 + (1 - w_1)\bar{x}_2$$
$$\bar{z}_2 = w_2 \bar{x}_1 + (1 - w_2)\bar{x}_2$$

$$\begin{aligned} SSZ_1 &= \sum z_{i1}^2 - \frac{(\sum z_{i1})^2}{n} = \sum (w_1 x_{i1} + (1 - w_1) x_{i2})^2 - n(w_1 \bar{x}_1 + (1 - w_1)\bar{x}_2)^2 \\ &= w_1^2 \sum x_{i1}^2 + (1 - w_1)^2 \sum x_{i2}^2 + 2w_1(1 - w_1) \sum x_{i1} x_{i2} \\ &\quad - n(w_1^2 \bar{x}_1^2 + (1 - w_1)^2 \bar{x}_2^2 + 2w_1(1 - w_1)\bar{x}_1 \bar{x}_2) \\ &= w_1^2 SSX_1 + (1 - w_1)^2 SSX_2 + 2w_1(1 - w_1) SSX_{12} \end{aligned}$$

Similarly for SSZ_2:

$$SSZ_2 = \sum z_{i2}^2 - \frac{(\sum z_{i2})^2}{n} = \cdots = w_2^2 SSX_1 + (1-w_2)^2 SSX_2 + 2w_2(1-w_1)SSX_{12}$$

$$SSZ_{12} = \sum z_{i1} z_{i2} - \frac{\sum z_{i1} \sum z_{i2}}{n} = \sum (z_{i1} - \bar{z}_1)(z_{i2} - \bar{z}_2)$$

$$= \sum [w_1 x_{i1} + (1-w_1)x_{i2} - w_1 \bar{x}_1 - (1-w_1)\bar{x}_2][w_2 x_{i1} + (1-w_2)x_{i2} - w_2 \bar{x}_1 - (1-w_2)\bar{x}_2]$$

$$= \sum [w_1(x_{i1} - \bar{x}_1) + (1-w_1)(x_{i2} - \bar{x}_2)][w_2(x_{i1} - \bar{x}_1) + (1-w_2)(x_{i2} - \bar{x}_2)]$$

$$= w_1 w_2 SSX_1 + (1-w_1)(1-w_2)SSX_2 + w_1(1-w_2)SSX_{12} + w_2(1-w_1)SSX_{12}$$

In the above variables,

$$SSX_1 = \sum x_{i1}^2 - \frac{(\sum x_{i1})^2}{n}, SSX_2 = \sum x_{i2}^2 - \frac{(\sum x_{i2})^2}{n}, SSX_{12} = \sum x_{i1} x_{i2} - \frac{\sum x_{i1} \sum x_{i2}}{n}, \bar{x}_1 = \frac{\sum x_{i1}}{n}, \bar{x}_2 = \frac{\sum x_{i2}}{n}.$$

Given the formulas for $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$, they will now be rewritten in terms of y, x_1, x_2. However, there are some common terms in the formulas that will first be written out. These terms include: $SSZ_1 \times SSZ_2 - SSZ_{12}^2$, $\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2$, $\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1$, and $-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1 \bar{z}_2 SSZ_{12}$.

First term:

$$\begin{aligned}
SSZ_1 \times SSZ_2 - SSZ_{12}^2 \\
&= [w_1^2 SSX_1 + (1-w_1)^2 SSX_2 + 2w_1(1-w_1)SSX_{12}][w_2^2 SSX_1 + (1-w_2)^2 SSX_2 \\
&\quad + 2w_2(1-w_1)SSX_{12}] \\
&\quad - [w_1 w_2 SSX_1 + (1-w_1)(1-w_2)SSX_2 + w_1(1-w_2)SSX_{12} \\
&\quad + w_2(1-w_1)SSX_{12}]^2 \\
&= w_1^2 w_2^2 SSX_1^2 + w_1^2(1-w_2)^2 SSX_1 SSX_2 + 2w_1^2 w_2(1-w_2)SSX_1 SSX_{12} \\
&\quad + (1-w_1)^2 w_2^2 SSX_1 SSX_2 + (1-w_1)^2(1-w_2)^2 SSX_2^2 \\
&\quad + 2(1-w_1)^2 w_2(1-w_2)SSX_2 SSX_{12} + 2w_1(1-w_1)w_2^2 SSX_1 SSX_{12} \\
&\quad + 2w_1(1-w_1)(1-w_2)^2 SSX_2 SSX_{12} + 4w_1 w_2(1-w_1)(1-w_2)SSX_{12}^2 \\
&\quad - [w_1^2 w_2^2 SSX_1^2 + (1-w_1)^2(1-w_2)^2 SSX_2^2 + w_1^2(1-w_2)^2 SSX_{12}^2 \\
&\quad + w_2^2(1-w_1)^2 SSX_{12}^2 + 2w_1 w_2(1-w_1)(1-w_2)SSX_1 SSX_2 \\
&\quad + 2w_1^2 w_2(1-w_2)SSX_1 SSX_{12} + 2w_1(1-w_1)w_2^2 SSX_1 SSX_{12} \\
&\quad + 2w_1(1-w_1)(1-w_2)^2 SSX_2 SSX_{12} + 2(1-w_1)^2 w_2(1-w_2)SSX_2 SSX_{12} \\
&\quad + 2w_1 w_2(1-w_1)(1-w_2)SSX_{12}^2] \\
&= w_1^2(1-w_2)^2 SSX_1 SSX_2 + (1-w_1)^2 w_2^2 SSX_1 SSX_2 + 2w_1 w_2(1-w_1)(1-w_2)SSX_{12}^2 \\
&\quad - w_1^2(1-w_2)^2 SSX_{12}^2 - w_2^2(1-w_1)^2 SSX_{12}^2 \\
&\quad - 2w_1 w_2(1-w_1)(1-w_2)^2 SSX_1 SSX_2 \\
&= (SSX_1 SSX_2 - SSX_{12}^2)[w_1^2(1-w_2)^2 + (1-w_1)^2 w_2^2 - 2w_1(1-w_2)w_2(1-w_1)] \\
&= (SSX_1 SSX_2 - SSX_{12}^2)[w_1(1-w_2) - w_2(1-w_2)]^2 \\
&= (SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2
\end{aligned}$$

Second term:

$$\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2$$
$$= (w_2\bar{x}_1 + (1-w_2)\bar{x}_2)[w_1 w_2 SSX_1 + (1-w_1)(1-w_2)SSX_2$$
$$+ w_1(1-w_2)SSX_{12} + w_2(1-w_1)SSX_{12}]$$
$$- (w_1\bar{x}_1 + (1-w_1)\bar{x}_2)[w_2^2 SSX_1 + (1-w_2)^2 SSX_2 + 2w_2(1-w_2)SSX_{12}]$$
$$= [w_1 w_2^2 \bar{x}_1 SSX_1 + (1-w_1)w_2(1-w_2)\bar{x}_1 SSX_2 + w_1 w_2(1-w_2)\bar{x}_1 SSX_{12}$$
$$+ (1-w_1)w_2^2\bar{x}_1 SSX_{12} + w_1 w_2(1-w_2)\bar{x}_2 SSX_{12} + (1-w_1)(1-w_2)\bar{x}_2 SSX_2$$
$$+ w_1(1-w_2)^2\bar{x}_2 SSX_{12} + (1-w_1)w_2(1-w_2)\bar{x}_2 SSX_{12}]$$
$$- [w_1 w_2^2\bar{x}_1 SSX_1 + w_1(1-w_2)^2\bar{x}_1 SSX_2 + 2w_1 w_2(1-w_2)\bar{x}_1 SSX_{12}$$
$$+ (1-w_1)w_2^2\bar{x}_2 SSX_1 + (1-w_1)(1-w_2)\bar{x}_2 SSX_2$$
$$+ 2(1-w_1)w_2(1-w_2)\bar{x}_2 SSX_{12}]$$
$$= [(1-w_1)w_2(1-w_2) - w_1(1-w_2)^2]\bar{x}_1 SSX_2 + [(1-w_1)w_2^2 - w_1 w_2(1-w_2)]\bar{x}_1 SSX_{12}$$
$$+ [w_1 w_2(1-w_2) - (1-w_1)w_2^2]\bar{x}_2 SSX_1$$
$$+ [w_1(1-w_2)^2 - (1-w_1)w_2(1-w_2)]\bar{x}_2 SSX_{12}$$
$$= (w_2 - w_1)(1-w_2)\bar{x}_1 SSX_2 + (w_2 - w_1)w_2\bar{x}_1 SSX_{12} + (w_1 - w_2)w_2\bar{x}_2 SSX_1$$
$$+ (w_1 - w_2)(1-w_2)\bar{x}_2 SSX_{12}$$
$$= (w_2 - w_1)[(1-w_2)\bar{x}_1 SSX_2 + w_2\bar{x}_1 SSX_{12} - w_2\bar{x}_2 SSX_1 - (1-w_2)\bar{x}_2 SSX_{12}]$$

Third term:

$$\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1$$
$$= (w_1\bar{x}_1 + (1-w_1)\bar{x}_2)[w_1 w_2 SSX_1 + (1-w_1)(1-w_2)SSX_2$$
$$+ w_1(1-w_2)SSX_{12} + w_2(1-w_1)SSX_{12}]$$
$$- (w_2\bar{x}_1 + (1-w_1)\bar{x}_2)[w_1^2 SSX_1 + (1-w_1)^2 SSX_2 + 2w_1(1-w_1)SSX_{12}]$$
$$= w_1^2 w_2\bar{x}_1 SSX_1 + w_1(1-w_1)(1-w_2)\bar{x}_1 SSX_2 + w_1^2(1-w_2)\bar{x}_1 SSX_{12}$$
$$+ w_1(1-w_1)w_2\bar{x}_1 SSX_{12} + w_1(1-w_1)w_2\bar{x}_2 SSX_1 + (1-w_1)^2(1-w_2)\bar{x}_2 SSX_2$$
$$+ w_1(1-w_1)(1-w_2)\bar{x}_2 SSX_{12} + (1-w_1)^2 w_2\bar{x}_2 SSX_{12}$$
$$- [w_1^2 w_2\bar{x}_1 SSX_1 + (1-w_1)^2 w_2\bar{x}_1 SSX_2 + 2w_1(1-w_1)w_2\bar{x}_1 SSX_{12}$$
$$+ w_1^2(1-w_2)\bar{x}_2 SSX_1 + (1-w_1)^2(1-w_2)\bar{x}_2 SSX_2$$
$$+ 2w_1(1-w_1)(1-w_2)\bar{x}_2 SSX_{12}]$$
$$= [w_1(1-w_2) - (1-w_1)w_2](1-w_1)\bar{x}_1 SSX_2 + [(1-w_1)w_2 - w_1(1-w_2)]w_1\bar{x}_2 SSX_1$$
$$+ [w_1(1-w_2) - (1-w_1)w_2]w_1\bar{x}_1 SSX_{12}$$
$$+ [(1-w_1)w_2 - w_1(1-w_2)](1-w_1)\bar{x}_2 SSX_{12}$$
$$= (w_1 - w_2)[(1-w_1)\bar{x}_1 SSX_2 - w_1\bar{x}_2 SSX_1 + w_1\bar{x}_1 SSX_{12} - (1-w_1)\bar{x}_2 SSX_{12}]$$

Last term:

$$-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1\bar{z}_2 SSZ_{12} = \bar{z}_1[\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2] + \bar{z}_2[\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1]$$
$$= \bar{z}_1 A + \bar{z}_2 B$$
$$A = \bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2$$
$$= (w_2\bar{x}_1 + (1-w_1)\bar{x}_2)[w_1 w_2 SSX_1 + (1-w_1)(1-w_2)SSX_2 + w_1(1-w_2)SSX_{12}$$
$$+ w_2(1-w_1)SSX_{12}]$$
$$- (w_1\bar{x}_1 + (1-w_1)\bar{x}_2)[w_2^2 SSX_1 + (1-w_2)^2 SSX_2 + 2w_2(1-w_2)SSX_{12}]$$

$$= w_1 w_2^2 \bar{x}_1 SSX_1 + (1-w_1) w_2 (1-w_2) \bar{x}_1 SSX_2 + w_1 w_2 (1-w_2) \bar{x}_1 SSX_{12}$$
$$+ (1-w_1) w_2^2 \bar{x}_1 SSX_{12} + w_1 w_2 (1-w_2) \bar{x}_2 SSX_1 + (1-w_1)(1-w_2)^2 \bar{x}_2 SSX_2$$
$$+ w_1 (1-w_2)^2 \bar{x}_2 SSX_{12} + (1-w_1) w_2 (1-w_2) \bar{x}_2 SSX_{12}$$
$$- [w_1 w_2^2 \bar{x}_1 SSX_1 + w_1 (1-w_2)^2 \bar{x}_1 SSX_2 + 2 w_1 w_2 (1-w_2) \bar{x}_1 SSX_{12}$$
$$+ (1-w_1) w_2^2 \bar{x}_2 SSX_1 + (1-w_1)(1-w_2)^2 \bar{x}_2 SSX_2$$
$$+ 2(1-w_1) w_2 (1-w_2) \bar{x}_2 SSX_{12}]$$
$$= [(1-w_1) w_2 - w_1 (1-w_2)](1-w_2) \bar{x}_1 SSX_2 + [w_1 (1-w_2) - (1-w_1) w_2] w_2 \bar{x}_2 SSX_1$$
$$+ [(1-w_1) w_2 - w_1 (1-w_2)] w_2 \bar{x}_1 SSX_{12}$$
$$+ [w_1 (1-w_2) - (1-w_1) w_2](1-w_2) \bar{x}_2 SSX_{12}$$
$$= (w_2 - w_1)[(1-w_2) \bar{x}_1 SSX_2 - w_2 \bar{x}_2 SSX_1 + w_2 \bar{x}_1 SSX_{12} - (1-w_2) \bar{x}_2 SSX_{12}]$$

$$B = \bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1$$
$$= (w_1 \bar{x}_1 + (1-w_1) \bar{x}_2)[w_1 w_2 SSX_1 + (1-w_1)(1-w_2) SSX_2 + w_1 (1-w_2) SSX_{12}$$
$$+ w_2 (1-w_1) SSX_{12}]$$
$$- (w_2 \bar{x}_1 + (1-w_2) \bar{x}_2)[w_1^2 SSX_1 + (1-w_1)^2 SSX_2 + 2 w_1 (1-w_1) SSX_{12}]$$
$$= w_1^2 w_2 \bar{x}_1 SSX_1 + w_1 (1-w_1)(1-w_2) \bar{x}_1 SSX_2 + w_1^2 (1-w_2) \bar{x}_1 SSX_{12}$$
$$+ w_1 (1-w_1) w_2 \bar{x}_1 SSX_{12} + w_1 (1-w_1) w_2 \bar{x}_2 SSX_1 + (1-w_1)^2 (1-w_2) \bar{x}_2 SSX_2$$
$$+ w_1 (1-w_1)(1-w_2) \bar{x}_2 SSX_{12} + (1-w_1)^2 w_2 \bar{x}_2 SSX_{12}$$
$$- [w_1^2 w_2 \bar{x}_1 SSX_1 + (1-w_1)^2 w_2 \bar{x}_1 SSX_2 + 2 w_1 (1-w_1) w_2 \bar{x}_1 SSX_{12}$$
$$+ w_1^2 (1-w_2) \bar{x}_2 SSX_1 + (1-w_1)^2 (1-w_2) SSX_2$$
$$+ 2 w_1 (1-w_1)(1-w_2) \bar{x}_2 SSX_{12}]$$
$$= [w_1 (1-w_2) - (1-w_1) w_2](1-w_1) \bar{x}_1 SSX_2 + [(1-w_1) w_2 - w_1 (1-w_2)] w_1 \bar{x}_2 SSX_1$$
$$+ [w_1 (1-w_2) - (1-w_1) w_2] w_1 \bar{x}_1 SSX_{12}$$
$$+ [(1-w_1) w_2 - w_1 (1-w_2)](1-w_1) \bar{x}_2 SSX_{12}$$
$$= (w_1 - w_2)[(1-w_1) \bar{x}_1 SSX_2 - w_1 \bar{x}_2 SSX_1 + w_1 \bar{x}_1 SSX_{12} - (1-w_1) \bar{x}_2 SSX_{12}]$$

$$\cdots = \bar{z}_1 A + \bar{z}_2 B$$
$$= \bar{z}_1 \{(w_2 - w_1)[(1-w_2) \bar{x}_1 SSX_2 - w_2 \bar{x}_2 SSX_1 + w_2 \bar{x}_1 SSX_{12} - (1-w_2) \bar{x}_2 SSX_{12}]\}$$
$$+ \bar{z}_2 \{(w_1 - w_2)[(1-w_1) \bar{x}_1 SSX_2 - w_1 \bar{x}_2 SSX_1 + w_1 \bar{x}_1 SSX_{12}$$
$$- (1-w_1) \bar{x}_2 SSX_{12}]\}$$
$$= (w_2 - w_1)\{(w_1 \bar{x}_1 + (1-w_1) \bar{x}_2)[(1-w_2) \bar{x}_1 SSX_2 - w_2 \bar{x}_2 SSX_1 + w_2 \bar{x}_1 SSX_{12}$$
$$- (1-w_2) \bar{x}_2 SSX_{12}]$$
$$- (w_2 \bar{x}_1 + (1-w_2) \bar{x}_2)[(1-w_1) \bar{x}_1 SSX_2 - w_1 \bar{x}_2 SSX_1 + w_1 \bar{x}_1 SSX_{12}$$
$$- (1-w_1) \bar{x}_2 SSX_{12}]\}$$
$$= (w_2 - w_1)\{w_1 (1-w_2) \bar{x}_1^2 SSX_2 - w_1 w_2 \bar{x}_1 \bar{x}_2 SSX_1 + w_1 w_2 \bar{x}_1^2 SSX_{12} - w_1 (1-w_2) \bar{x}_1 \bar{x}_2 SSX_{12}$$
$$+ (1-w_1)(1-w_2) \bar{x}_1 \bar{x}_2 SSX_2 - (1-w_1) w_2 \bar{x}_2^2 SSX_1 + (1-w_1) w_2 \bar{x}_1 \bar{x}_2 SSX_{12}$$
$$- (1-w_1)(1-w_2) \bar{x}_2^2 SSX_{12} - (1-w_1) w_2 \bar{x}_1^2 SSX_2 + w_1 w_2 \bar{x}_1 \bar{x}_2 SSX_1$$
$$- w_1 w_2 \bar{x}_1^2 SSX_{12} + (1-w_1) w_2 \bar{x}_1 \bar{x}_2 SSX_{12} - (1-w_1)(1-w_2) \bar{x}_1 \bar{x}_2 SSX_2$$
$$+ w_1 (1-w_2) \bar{x}_2^2 SSX_1 - w_1 (1-w_2) \bar{x}_1 \bar{x}_2 SSX_{12} + (1-w_1)(1-w_2) \bar{x}_2^2 SSX_{12}\}$$
$$= (w_2 - w_1)\{[w_1 (1-w_2) - (1-w_1) w_2] \bar{x}_1^2 SSX_2 + [w_1 (1-w_2) - (1-w_1) w_2] \bar{x}_2^2 SSX_1$$
$$+ [(1-w_1) w_2 - w_1 (1-w_2)] 2 \bar{x}_1 \bar{x}_2 SSX_{12}\}$$
$$= (w_2 - w_1)\{(w_1 - w_2) \bar{x}_1^2 SSX_2 + (w_1 - w_2) \bar{x}_2^2 SSX_1 + (w_2 - w_1) 2 \bar{x}_1 \bar{x}_2 SSX_{12}\}$$
$$= -(w_2 - w_1)^2 (\bar{x}_1^2 SSX_2 + \bar{x}_2^2 SSX_1 - 2 \bar{x}_1 \bar{x}_2 SSX_{12})$$

Therefore the $\hat{\beta}_j$ terms can be rewritten as follows:

$$\hat{\beta}_0 = \frac{\sum y_i}{n} - \frac{-\bar{z}_2^2 SSZ_1 - \bar{z}_1^2 SSZ_2 + 2\bar{z}_1\bar{z}_2 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum y_i + \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i1}\, y_i + \frac{\bar{z}_1 SSZ_{12} - \bar{z}_2 SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i2}\, y_i$$

$$= \frac{\sum y_i}{n} - \frac{-(w_2 - w_1)^2(\bar{x}_1^2 SSX_2 + \bar{x}_2^2 SSX_1 - 2\bar{x}_1\bar{x}_2 SSX_{12})}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i + \frac{(w_2 - w_1)[(1 - w_2)\bar{x}_1 SSX_2 + w_2\bar{x}_1 SSX_{12} - w_2\bar{x}_2 SSX_1 - (1 - w_2)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i + \frac{(w_1 - w_2)[(1 - w_1)\bar{x}_1 SSX_2 - w_1\bar{x}_2 SSX_1 + w_1\bar{x}_1 SSX_{12} - (1 - w_1)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

$$\hat{\beta}_1 = \frac{\bar{z}_2 SSZ_{12} - \bar{z}_1 SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum y_i + \frac{SSZ_2}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i1}\, y_i - \frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i2}\, y_i$$

$$= \frac{(w_2 - w_1)[(1 - w_2)\bar{x}_1 SSX_2 + w_2\bar{x}_1 SSX_{12} - w_2\bar{x}_2 SSX_1 - (1 - w_2)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i + \frac{w_2^2 SSX_1 + (1 - w_2)^2 SSX_2 + 2w_2(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i - \frac{w_1 w_2 SSX_1 + (1 - w_1)(1 - w_2)SSX_2 + w_1(1 - w_2)SSX_{12} + w_2(1 - w_1)SSX_{12}}{S(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

$$\hat{\beta}_2 = \frac{-\bar{z}_2 SSZ_1 + \bar{z}_1 SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum y_i - \frac{SSZ_{12}}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i1}\, y_i + \frac{SSZ_1}{SSZ_1 \times SSZ_2 - SSZ_{12}^2}\sum z_{i2}\, y_i$$

$$= \frac{(w_1 - w_2)[(1 - w_1)\bar{x}_1 SSX_2 - w_1\bar{x}_2 SSX_1 + w_1\bar{x}_1 SSX_{12} - (1 - w_1)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i - \frac{w_1 w_2 SSX_1 + (1 - w_1)(1 - w_2)SSX_2 + w_1(1 - w_2)SSX_{12} + w_2(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i + \frac{w_1^2 SSX_1 + (1 - w_1)^2 SSX_2 + 2w_1(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum [w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

So, it has been shown that:

$$\hat{\beta}_0 = \frac{\sum y_i}{n} - \frac{-(w_2 - w_1)^2(\bar{x}_1^2 SSX_2 + \bar{x}_2^2 SSX_1 - 2\bar{x}_1\bar{x}_2 SSX_{12})}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i + \frac{(w_2 - w_1)[(1 - w_2)\bar{x}_1 SSX_2 + w_2\bar{x}_1 SSX_{12} - w_2\bar{x}_2 SSX_1 - (1 - w_2)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i + \frac{(w_1 - w_2)[(1 - w_1)\bar{x}_1 SSX_2 - w_1\bar{x}_2 SSX_1 + w_1\bar{x}_1 SSX_{12} - (1 - w_1)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

$$\hat{\beta}_1 = \frac{(w_2 - w_1)[(1 - w_2)\bar{x}_1 SSX_2 + w_2\bar{x}_1 SSX_{12} - w_2\bar{x}_2 SSX_1 - (1 - w_2)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i + \frac{w_2^2 SSX_1 + (1 - w_2)^2 SSX_2 + 2w_2(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i - \frac{w_1 w_2 SSX_1 + (1 - w_1)(1 - w_2)SSX_2 + w_1(1 - w_2)SSX_{12} + w_2(1 - w_1)SSX_{12}}{S(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

$$\hat{\beta}_2 = \frac{(w_1 - w_2)[(1 - w_1)\bar{x}_1 SSX_2 - w_1\bar{x}_2 SSX_1 + w_1\bar{x}_1 SSX_{12} - (1 - w_1)\bar{x}_2 SSX_{12}]}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum y_i - \frac{w_1 w_2 SSX_1 + (1 - w_1)(1 - w_2)SSX_2 + w_1(1 - w_2)SSX_{12} + w_2(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_1 x_{i1} + (1 - w_1)x_{i2}]\, y_i + \frac{w_1^2 SSX_1 + (1 - w_1)^2 SSX_2 + 2w_1(1 - w_1)SSX_{12}}{(SSX_1 SSX_2 - SSX_{12}^2)(w_1 - w_2)^2}\sum[w_2 x_{i1} + (1 - w_2)x_{i2}]\, y_i$$

where $SSX_1 = \sum x_{i1}^2 - \frac{(\sum x_{i1})^2}{n}, SSX_2 = \sum x_{i2}^2 - \frac{(\sum x_{i2})^2}{n}, SSX_{12} = \sum x_{i1}x_{i2} - \frac{\sum x_{i1}\sum x_{i2}}{n}, \bar{x}_1 = \frac{\sum x_{i1}}{n}, \bar{x}_2 = \frac{\sum x_{i2}}{n}$.

b. Derive a statistical test for testing statistical significance of the regression model for (y, x_1, x_2), not for (y, z_1, z_2).

<u>Ans:</u> (*Note: Answer partially used from Problem 1.*)

A hypothesis test that is checking if there is a linear relationship between the response y and any of the regressor variables z_1, z_2 appear as follows,

$$H_0: \beta_1 = \beta_2 = 0 \text{ vs. } H_1: \beta_j \neq 0 \text{ for at least one } j \text{ for } j \in \{1,2\}.$$

The test statistic for the above hypothesis, F_0, appears as follows,

$$F_0 = \frac{\frac{SS_R}{k}}{\frac{SS_{Res}}{n-k-1}} = \frac{MS_R}{MS_{Res}},$$

where

$$MS_R = \frac{SS_R}{k} = \frac{1}{k}\left[\widehat{\boldsymbol{\beta}}'\mathbf{Z}'\mathbf{y} - \frac{(\sum_{i=1}^{n} y_i)^2}{n}\right],$$
$$MS_{Res} = \frac{SS_{Res}}{n-k-1} = \frac{1}{n-k-1}\left[\mathbf{y}'\mathbf{y} - \widehat{\boldsymbol{\beta}}'\mathbf{Z}'\mathbf{y}\right],$$
$$\widehat{\boldsymbol{\beta}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}.$$

In the above, it is also the case that there is a total of n observations, for y_i, $i = 1, \cdots, n$. Another assumption is that $(\mathbf{Z}'\mathbf{Z})^{-1}$ exists, in other words $\mathbf{Z}'\mathbf{Z}$ is invertible. This requires that the columns of $\mathbf{Z}$ are linearly independent. Here, Z is the design matrix that includes the variables z_1 and z_2 in addition to a column of 1's to match with the intercept term β_0.

The problem however is asking that the test statistic be written in terms of x rather than z. Therefore, some changes need to be made to the variables for MS_R and MS_{Res}. It has already been shown in part a) how $\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$ can be written in terms of x rather than z. Therefore, that step won't be repeated. Also, the formulas for MS_R and MS_{Res} includes a few more variables such as: n, k, $\mathbf{y}$, and $\mathbf{Z}$. The first three terms, n, k, and $\mathbf{y}$, do not depend on z, and so they remain unchanged for this part of the problem. However, $\mathbf{Z}$ needs to be expressed in terms of x's rather than z's. Denote a new matrix, $\mathbf{X}^*$, as follows,

$$\mathbf{X}^* = [\mathbf{1} \quad \mathbf{x}_1^{*\prime} \quad \mathbf{x}_2^{*\prime}] = \begin{bmatrix} 1 & w_1x_{11} + (1-w_1)x_{12} & w_2x_{11} + (1-w_2)x_{12} \\ 1 & w_1x_{21} + (1-w_1)x_{22} & w_2x_{21} + (1-w_2)x_{22} \\ \vdots & \vdots & \vdots \\ 1 & w_1x_{n1} + (1-w_1)x_{n2} & w_2x_{n1} + (1-w_2)x_{n2} \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & z_{13} \\ 1 & z_{22} & z_{23} \\ \vdots & \vdots & \vdots \\ 1 & z_{n2} & z_{n3} \end{bmatrix}$$
$$= \mathbf{Z}.$$

Now, the test statistic, F_0, can be written as follows,

$$F_0 = \frac{\frac{SS_R}{k}}{\frac{SS_{Res}}{n-k-1}} = \frac{MS_R}{MS_{Res}} = \frac{(n-3)}{2} \cdot \frac{\left[\widehat{\boldsymbol{\beta}}'\mathbf{X}^{*\prime}\mathbf{y} - \frac{(\sum_{i=1}^{n} y_i)^2}{n}\right]}{\left[\mathbf{y}'\mathbf{y} - \widehat{\boldsymbol{\beta}}'\mathbf{X}^{*\prime}\mathbf{y}\right]}.$$

To test the hypothesis, then the null can be rejected if $F_0 > F_{\alpha,k,n-k-1}$, where α is the confidence level for the hypothesis test and F is the F-distribution with degrees of freedom k and $n-k-1$. If we reject the null, then we can say that at least one of the regressors $\mathbf{x}_1^*$ or $\mathbf{x}_2^*$ (i.e., z_1 or z_2) contributes significantly to the model.

State the assumptions for each step of your discussion or derivation in a), b), c).

4. Suppose that n subjects give data following the true model
$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \varepsilon,$$
where x_1 and x_2 are non-random regressors, and ε is the random error.

   Now suppose that the reduced model
$$y = \beta_0 + \beta_1x_1 + \varepsilon,$$
is also fitted to the same data to obtain the least-squares estimator $\tilde{\beta}_1$.

Discuss by mathematical arguments whether this least-squares estimator $\tilde{\beta}_1$ is biased for β_1. If yes, discuss by mathematical arguments the conditions under which $\tilde{\beta}_1$ is unbiased for β_1.

<u>Ans:</u>

To show whether or not $\tilde{\beta}_1$ is unbiased, it must be shown that $E(\tilde{\beta}_1) = \beta_1$. Using the simple linear regression model, denote the OLS estimators of the reduced model with $\tilde{\beta}_0$ and $\tilde{\beta}_1$. Then, these estimators can be shown to be

$$\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1 \bar{x} \text{ and } \tilde{\beta}_1 = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)(y_i - \bar{y})}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = \sum_{i=1}^{n} c_i y_i,$$

where $c_i = \frac{x_{i1} - \bar{x}_1}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2}$. This is done by optimizing the following formula,

$$S(\beta) = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_{i1})^2.$$

Note: Let x_{ij} represent the i^{th} sample from the j^{th} variable, where $j \in \{1,2\}$. It has been shown previously in problem 2 b) how $\sum_{i=1}^{n} c_i y_i$ can be derived.

Now, let us analyze the expected value of $\tilde{\beta}_1$:

$$E(\tilde{\beta}_1) = E\left(\sum_{i=1}^{n} c_i y_i\right) = \sum_{i=1}^{n} c_i E(y_i)$$

Note: The assumption above is that x is non-random and so it can be treated as a constant term for the expectation.

$$\cdots = \sum_{i=1}^{n} c_i E(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i) = \sum_{i=1}^{n} c_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})$$

Note: An assumption used above is that $E(\varepsilon_i) = 0$.

$$\cdots = \beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} c_i x_{i1} + \beta_2 \sum_{i=1}^{n} c_i x_{i2} = \beta_1 + \beta_2 \sum_{i=1}^{n} c_i x_{i2} = \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)x_{i2}}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2}$$

Note: An assumption used above is that $\sum_{i=1}^{n} c_i = 0$, $\sum_{i=1}^{n} c_i x_{i1} = 1$, and that $\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2$ is nonzero. The first two will be shown briefly.

$$\sum_{i=1}^{n} c_i = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = \frac{\sum_{i=1}^{n} x_{i1} - \sum_{i=1}^{n} \bar{x}_1}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = \frac{n\bar{x}_1 - n\bar{x}_1}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = 0$$

$$\sum_{i=1}^{n} c_i x_{i2} = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)x_{i1}}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2} = 1$$

So, from $E(\tilde{\beta}_1) = \beta_1 + \beta_2 \frac{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)x_{i2}}{\sum_{j=1}^{n}(x_{j1} - \bar{x}_1)^2}$, it can clearly be seen that $E(\tilde{\beta}_1) \neq \beta_1$, so $\tilde{\beta}_1$ is biased.

There are special cases where $E(\tilde{\beta}_1) = \beta_1$ however. It can be seen that if either β_2 or

$\frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)x_{i2}}{\sum_{j=1}^{n}(x_{j1}-\bar{x}_1)^2}$ evaluate to 0, then $E\left(\tilde{\beta}_1\right) = \beta_1$. An example would be if x_1 and x_2 were identical, or some other way where the numerator $\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)x_{i2} = 0$.

State the assumptions for each step of your discussion or derivation.