Worksheet: Analysis of Variance for Regression

<table>
<thead>
<tr>
<th>AB_{11}</th>
<th>AB_{12}</th>
<th>AB_{13}</th>
<th>B_{1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB_{21}</td>
<td>AB_{22}</td>
<td>AB_{23}</td>
<td>B_{2}</td>
</tr>
<tr>
<td>A_{1}</td>
<td>A_{2}</td>
<td>A_{3}</td>
<td></td>
</tr>
</tbody>
</table>

\[ s_x = ( \quad ) \rightarrow X \quad s_y = ( \quad ) \rightarrow Y \quad r = ( \quad ) \rightarrow R \]

**Regression line** \( y = a + bx \) goes through the mean point \((\bar{x}, \bar{y})\)

Your calculator should directly calculate \( a = ( \quad ) \rightarrow A \) and \( b = ( \quad ) \rightarrow B \).

If not, then \( b = \frac{rs_y}{s_x} \) and \( a = \bar{y} - b\bar{x} \) since \( \bar{y} = a + b\bar{x} \).

\[ s_{xy} = rs_x s_y = ( \quad ) \rightarrow Z \]

\[ SS_x = S(x_i - \bar{x})^2 = (n-1)s_x^2 = ( \quad ) \rightarrow X \]

\[ SS_y = S(y_i - \bar{y})^2 = (n-1)s_y^2 = ( \quad ) \rightarrow Y \]

\[ SS_{xy} = S(x_i - \bar{x})(y_i - \bar{y}) = (n-1)s_{xy} = ( \quad ) \rightarrow Z \]

\[ SSR = r^2 SS_y = ( \quad ) \rightarrow Z \]

\[ SSE = (1 - r^2)SS_y \]

\[ df_R = 1, \quad MSR = SSR/df_R = SSR/1 = SSR = R \]

\[ df_e = n - 2, \quad MSE = SSE/df_e = E/df_e = (\quad) \rightarrow E \]

\[ s = \sqrt{MSE} \]

**Test for significant linear relationship between X and Y.**

\( H_0: \) There is. \( H_0: \) There is no significant linear relation, \( MSR < MSE, \ F < 1 \) where the test statistic is

\( F = \frac{MSR}{MSE} = \frac{(n - 2)r^2}{(1 - r^2)} = \frac{R}{E} \) (calculate \( F \) with both formulas as a check).

\[ df_1 = df_R = 1, \quad df_2 = df_e = n - 2. \quad Acceptance \ region: [0, F_{a}]. \]

**Theorem.** \( r^2 \) is proportion of variation due to regression. \( r^2 \) is called the **coefficient of determination.**

**Confidence interval for slope \( b \) when estimated with \( b. \)** \( t_{a/2} = ( \quad ) \rightarrow T \)

Standard error for \( b: \) \( SE = \sqrt{\frac{MSE}{SS_x}} = \sqrt{\frac{E}{X}} = ( \quad ) \rightarrow S. \) Confidence interval for \( b: \) \( b \pm t_{a/2}SE = B! TS. \)

**Given \( x = x_0, \) the confidence interval for the true or mean value \( y = a + bx_0 \) when estimated by \( \hat{y} = a + bx_0. \)**

Standard error for \( \hat{y}: \) \( SE = \sqrt{MSE(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_x})} = ( \quad ) \rightarrow S. \) Confidence interval for \( y = a + bx_0: \) \( (a + bx_0) \pm t_{a/2}SE. \)

**Given \( x = x_0, \) the confidence interval for the measured value \( y_0 = a + bx_0 + e \) when estimated by \( \hat{y} = a + bx_0. \)**

Standard error for \( \hat{y}: \) \( SE = \sqrt{MSE(1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_x})} = ( \quad ) \rightarrow S. \) Confidence interval for \( y_0: \) \( (a + bx_0) \pm t_{a/2}SE. \)