Because the samples are independent, the sample variances $S_i^2$’s are independent; therefore, $SSW/\sigma^2 \sim \chi^2_{I(J-1)}$.

Note
(1) Part I implies that $E(\text{SSW}) = I(J-1)\sigma^2$.
(2) $S_p^2 = MSW = \frac{SSW}{I(J-1)}$ is called the pooled sample variance.

**Proof of B.2**
Consider the sample with the sample means:

\[ \{\bar{Y}_1, \bar{Y}_2, \cdots, \bar{Y}_I\} \]

Since (1) each of them is a linear combination of independent normal random variables and (2) they are calculated from independent samples, they are independent normal random variables. In fact, $\bar{Y}_i \overset{\text{independent}}{\sim} N(\mu + \alpha_i, \sigma^2/J)$. Under the null hypothesis, they have the same population mean and variance. Thus, we can treat

\[ \{\bar{Y}_1, \bar{Y}_2, \cdots, \bar{Y}_I\} \]

as a random sample from $N(\mu, \sigma^2/I)$ when the null hypothesis is true.

The corresponding sample mean and sample variances are $\bar{Y} = \frac{\bar{Y}_1 + \cdots + \bar{Y}_I}{I}$ and $S_{TR}^2 = \frac{1}{I-1} \sum_{i=1}^{I} (\bar{Y}_i - \bar{Y})^2$, respectively. Based on the properties of sample variance (see stat120B), we have

\[
\frac{(I-1)S_{TR}^2}{\sigma^2/J} \stackrel{H_0}{\sim} \chi^2_{I-1}
\]

Now return to $SSB$.

\[
SSB/\sigma^2 = \frac{J}{\sigma^2} \sum_{i=1}^{I} (\bar{Y}_i - \bar{Y})^2
\]

\[
= \frac{(I-1)S_{TR}^2}{\sigma^2/J} \stackrel{H_0}{\sim} \chi^2_{I-1}
\]

**Proof of B.3**
$SSW$ is a function of $S_i^2, i = 1, \cdots, I$, where $S_i^2 = \frac{1}{J} \sum_{j=1}^{J} (Y_{ij} - \bar{Y}_i)^2$.

$SSB$ is a function of $\bar{Y}_i$’s ($\bar{Y}$ is also a function of $\bar{Y}_i$’s).

We claim that $\bar{Y}_1, \cdots, \bar{Y}_I$ and $S_1^2, \cdots, S_I^2$ are independent with each other.

When $i \neq i'$, $S_i^2$ and $\bar{Y}_{i'}$ are independent because they are functions of different observations.
When \( i = i' \), by 120statB, \( S_i^2 \) and \( \bar{Y}_{i'} \) are independent. (for a normal random sample, the sample mean and sample variance are independent)

Since \( SSB \) is a function of \( \bar{Y}_1, \ldots, \bar{Y}_I \), and \( SSW \) is a function of \( S_1^2, \ldots, S_I^2 \), \( SSB \) and \( SSW \) are independent from each other.

**Summary of the proof:**

to prove B.1, we consider samples \( \{Y_{i1}, Y_{i2}, \ldots, Y_{ij}\} \) for each \( i \);
to prove B.2, we consider the sample of means \( \{\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_I\} \).

Finally, we have all the elements for **Theorem C**

**Theorem C: The F-test for the One-Way ANOVA** Assume that the assumptions of one-way ANOVA hold (normality, independence, and equal variance), when the null hypothesis \( H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_I = 0 \) (or \( H_0 : \mu_1 = \mu_2 = \cdots = \mu_I \)) is true,

\[
F = \frac{MSB}{MSW} = \frac{SSB/(I-1)}{SSW/(I(J-1))} \sim F_{I-1,I(J-1)}
\]

**Proof:**

\[
F = \frac{MSB}{MSW} = \frac{SSB/(I-1)}{SSW/(I(J-1))} = \frac{SSB}{\sigma^2/(I-1)} \frac{SSW}{\sigma^2/(I(J-1))}
\]

Since \( SSW / \sigma^2 \sim \chi^2_{I(J-1)} \) and \( SSB / \sigma^2 \sim \chi^2_{I-1} \) and \( SSW \) and \( SSB \) are independent, \( F \sim F_{I-1,I(J-1)} \).

We reject \( H_0 \) at a significance level \( \alpha \) if the test statistic \( F \) is greater than \( F_{I-1,I(J-1),1-\alpha} \). Here \( F_{I-1,I(J-1),1-\alpha} \) is the upper \( \alpha \) point of \( F_{I-1,I(J-1)} \).

**The ANOVA table:**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>( SSB = \sum_{i=1}^I J(Y_{i1} - \bar{Y}_{i \cdot})^2 )</td>
<td>( I - 1 )</td>
<td>( MSB = \frac{SSB}{I-1} )</td>
<td>( MSB/MSW )</td>
</tr>
<tr>
<td>Error</td>
<td>( SSW = \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i \cdot})^2 )</td>
<td>( I(J-1) )</td>
<td>( MSW = SSW/(I(J-1)) )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( SSTO = \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{\cdot \cdot})^2 )</td>
<td>( IJ - 1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA table for the example
\[
\begin{array}{c|ccccc}
\text{Source} & SS & df & MS & F \\
\hline
\text{Labs} & .125 & 6 & .021 & 5.66 \\
\text{Error} & .231 & 63 & .0037 & \\
\text{Total} & .356 & 69 & & \\
\end{array}
\]

\( F = 5.66 > F_{6.63.0.95} = 2.246408 \). And the p-value is much smaller than 0.01. So we conclude that the levels of the chemical compound are different across different labs.

Another useful result based on **Theorem B**

\[
\frac{(\bar{Y}_{i_1} - \bar{Y}_{i_2}) - (\alpha_{i_1} - \alpha_{i_2})}{S_p \sqrt{\frac{1}{J} + \frac{1}{J}}} \sim t_{I(J-1)}
\]

for \( i_1 \neq i_2 \). Here \( S_p^2 = MSW \).

Proof: \( \bar{Y}_{i_1} \sim N(\mu + \alpha_{i_1}, \sigma^2/J), \bar{Y}_{i_2} \sim N(\mu + \alpha_{i_2}, \sigma^2/J) \).

For different \( i_1, i_2, \bar{Y}_{i_1} \) and \( \bar{Y}_{i_2} \) are also independent. Therefore

\[
\frac{(\bar{Y}_{i_1} - \bar{Y}_{i_2}) - (\alpha_{i_1} - \alpha_{i_2})}{S_p \sqrt{\frac{1}{J} + \frac{1}{J}}} \sim N(0, 1)
\]

Based on B.2, we have \( SSW/\sigma^2 = I(J-1)MSW/\sigma^2 = I(J-1)S_p^2/\sigma^2 \sim \chi^2_{I(J-1)} \).

Also, in the proof of B.3, we have shown that the two vectors are independent. So \( \bar{Y}_{i_1} - \bar{Y}_{i_2} \) are independent of \( S_p^2 \).

Based on the above facts, we have

\[
\frac{(\bar{Y}_{i_1} - \bar{Y}_{i_2}) - (\alpha_{i_1} - \alpha_{i_2})}{\sqrt{I(J-1)S_p^2/\sigma^2I(J-1)}} \sim t_{I(J-1)}
\]

Simplify the left hand side,

\[
\frac{(\bar{Y}_{i_1} - \bar{Y}_{i_2}) - (\alpha_{i_1} - \alpha_{i_2})}{S_p \sqrt{\frac{1}{J} + \frac{1}{J}}} \sim t_{I(J-1)}
\]

**A special case:**

When \( I = 2 \), the two-sample t-test statistic \( t \sim t_{J-1} \). In the homework you have shown that \( F = t^2 \sim F_{1,J-1} \). This agrees with the fact that \( Z \sim t_n \Rightarrow Z^2 \sim F_{1,n} \).