Emma Erdtmann
Daniel Detore
I pledge my honor that I have abided by the Stevens Honor System.

Query 2

Problem 1

$$\bar{X}_{\cdot,\cdot} = \frac{1}{n_1+\cdots+n_k}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{i,j} = \frac{n_1\bar{X}_{1,\cdot}+\cdots+n_k\bar{X}_{k,\cdot}}{n_1+\cdots+n_k}.$$

From Slides 7:

$$\bar{x}_{\cdot,\cdot} = \cancel{\tfrac{1}{n}} \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{i,j} = \cancel{\tfrac{1}{n}} \sum_{i=1}^{k} n_i \bar{x}_{i,\cdot}$$

$$\Rightarrow \underbrace{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{i,j} = \sum_{i=1}^{k} n_i \bar{x}_{i,\cdot}}$$

$$\cancel{\frac{1}{n_1+\cdots+n_k}} \cdot \left[\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{i,j}\right] = \cancel{\frac{1}{n_1+\cdots+n_k}} \cdot \left[\frac{n_1\bar{x}_{1,\cdot}+\cdots+n_k\bar{x}_{k,\cdot}}{1}\right]$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{i,j} = n_1\bar{x}_{1,\cdot}+\cdots+n_k\bar{x}_{k,\cdot}$$

$$\sum_{i=1}^{k} n_i \bar{x}_{i,\cdot} = n_1\bar{x}_{1,\cdot}+\cdots+n_k\bar{x}_{k,\cdot}$$

$$n_1\bar{x}_{1,\cdot}+\dots n_k\bar{x}_{k,\cdot} = n_1\bar{x}_{1,\cdot}+\cdots+n_k\bar{x}_{k,\cdot}$$

Problem 2

$$SSB = \sum_{i=1}^{\overbrace{k}^{\#\text{ groups}}}\sum_{j=1}^{n_i}(\bar{X}_{i,\cdot}-\bar{X}_{\cdot,\cdot})^2 = \sum_{i=1}^{k} n_i(\bar{X}_{i,\cdot}-\bar{X}_{\cdot,\cdot})^2.$$

in each group, the # of observations is constant

Each group's sum of squares is just

$$\sum_{j=1}^{n_i}(\bar{x}_{i,\cdot}-\bar{x}_{\cdot,\cdot})^2 = n_i(\bar{x}_{i,\cdot}-\bar{x}_{\cdot,\cdot})^2$$

To get the sum of squares between all k groups we get the summation using this

$$\sum_{i=1}^{k} n_i(\bar{x}_{i,\cdot}-\bar{x}_{\cdot,\cdot})^2$$

Problem 3

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{i,j} - \bar{X}_{i,\cdot})^2 = \sum_{i=1}^{k} (n_i - 1) S_i^2.$$

From slides 7 (pg 20), Group sample variance

$$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{i,j} - \bar{x}_{i,\cdot})^2$$

we can cross multiply and see that

$$S_i^2 (n_i - 1) = \sum_{j=1}^{n_i} (x_{i,j} - \bar{x}_{i,\cdot})^2$$

So we can make this substitution into the LHS:

$$\boxed{SSE = \sum_{i=1}^{k} (n_i - 1) S_i^2}$$

Problem 4

According to lecture 7, $\psi = a_1\mu_1 + \ldots + a_k\mu_k$.

Problem 5

The lecture 7 notes give us the following: $\psi = a_1\mu_1 + \ldots + a_k\mu_k$, which we can rewrite as

$\psi = \sum_{i=1}^{k} a_i\mu_i$. The notes also tell us $C_\psi = \sum_{i=1}^{k} a_i\bar{X}_{i,\cdot}$, and we can approximate ψ as $E[C_\psi]$. We can simplify this knowing that $\bar{X}_{i,\cdot} \sim N(\mu_i, \sigma^2/n_i)$, $i = 1, \ldots, k$ are independent.

$$E[C_\psi] = E[\sum_{i=1}^{k} a_i\bar{X}_{i,\cdot}] = \sum_{i=1}^{k} a_i E[\bar{X}_{i,\cdot}] = \sum_{i=1}^{k} a_i\mu_i = \psi.$$

Problem 6
According to lecture 7:

$$Var[C_\psi] = Var[\sum_{i=1}^{k} a_i \overline{X}_{i,\cdot}] = \sum_{i=1}^{k} Var[a_i \overline{X}_{i,\cdot}] = \sum_{i=1}^{k} a_i^2 Var[\overline{X}_{i,\cdot}] = \sigma^2 \sum_{i=1}^{k} a_i^2/n \equiv \sigma^2_\psi.$$

This is possible because $\overline{X}_{i,\cdot} \sim N(\mu_i, \sigma^2/n_i)$, $i = 1, \ldots, k$ are independent.

Problem 7
C_ψ is a linear combination of standard normal random variables, therefore $C_\psi \sim N(\psi, \sigma^2_\psi)$

Problem 8
i
Under $H_0: \psi = 0$, C_ψ follows a standard normal distribution. Therefore
$P_{H_0}(|C_\psi| > 3) = 2(1 - pnorm(3)) = 0.002699796.$

```
> 2*(1-pnorm(3))
[1] 0.002699796
```

ii
Under $H_a: \psi > 0$, we must assume that $C_\psi > 0$. This is because C_ψ follows a normal distribution, so it is more likely that, even if we subtract 0.5, $C_\psi > 0$. Therefore
$P(C_\psi > 0) - \frac{1}{2} > 0$, thus its sign is positive.