Administrative

1. HW 1 due Today
2. Lab 2 on Wednesday
3. Lab 2 due on Thursday

Today

1. Multiple comparisons of pairs of populations means
2. The one-way effects model (PNC 9.2)
3. Analyzing data from randomized complete block experiments
Multiple comparisons

\[ H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \]
\[ H_a : \text{Not all the } \mu_i \text{ are equal.} \]

What if we reject the \( H_0 \) in favor of \( H_a \)?

- There is evidence of a difference between at least one pair of population means.
- There is evidence that \( \mu_i \neq \mu_j \) for at least one \( i \) and one \( j \neq i \).

How do we determine one(s)?

- We could test

  \[
  H_0 : \mu_i = \mu_j \\
  H_a : \mu_i \neq \mu_j
  \]

  for all possible pairs \((\mu_i, \mu_j)\).

- This is a situation where we want to compare the means of two independent normal populations with equal variances.
- We know how to do that!
1. Compute

\[ t^* = \frac{\bar{Y}_i - \bar{Y}_j}{\hat{\sigma} \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \]

- Chapter 6: When we only have two populations, \( S^2_p \) is our estimate of \( \sigma^2 \) based on the data from these two populations.
- Chapter 9: When we have more than two populations, the MSE is our estimate of \( \sigma^2 \) based on the data from all \( k \) populations.

2. When \( H_0 \) is true, \( t^* \sim t_{df} \).
   - Chapter 6: When we only have two populations, the degrees of freedom associated with estimating \( \sigma \) is \( n_1 + n_2 - 2 \).
   - Chapter 9: When we have more than two populations, the degrees of freedom associated with estimating \( \sigma \) is
     \[ df_E = n - k = (n_1 + n_2 + \cdots + n_k) - k. \]

3. If \( t^* \) is large in magnitude we reject \( H_0 \).
   - Chapter 6: When we only have two populations, we reject \( H_0 \) if
     \[ |t^*| > t_{df,1-\alpha/2}. \]
   - Chapter 9: Can we do the same thing?
• Recall that what we want to do is test

\[ H_0 : \mu_i = \mu_j \]
\[ H_a : \mu_i \neq \mu_j \]

for all possible pairs \((\mu_i, \mu_j)\). That’s \(k(k - 1)/2\) tests!

• Recall our discussion of the Type I error rate (i.e., \(\alpha\))

  - If we let \(\alpha = 0.05\), the probability that we incorrectly reject \(H_0\) for a single test is 0.05.
  - If we let \(\alpha = 0.05\) what is the probability that we incorrectly reject \(H_0\) for at least one out of our \(k(k - 1)/2\) tests?
  - If all the tests are independent:

\[
P(\text{at least one error}) = 1 - P(\text{no errors in } k(k - 1)/2 \text{ tests})
= 1 - (1 - \alpha)^{k(k-1)/2}
\]

<table>
<thead>
<tr>
<th>(k)</th>
<th>(k(k - 1)/2)</th>
<th>(1 - (1 - 0.05)^{k(k-1)/2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.0500000</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.1426250</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0.2649081</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.4012631</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>0.5367088</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>0.6594384</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>0.7621731</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>0.8422208</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>0.9005597</td>
</tr>
</tbody>
</table>

• What can we do to control this error rate?
  We will discuss two methods.
Bonferroni comparisons

- The Bonferroni method treats $\alpha$ as a Type I error “budget.” It then divides the up the error budget equally among the $k(k-1)/2$ tests being performed.

- So if we are performing $k(k-1)/2$ tests, use $\alpha^* = \frac{\alpha}{k(k-1)/2}$ for each individual test.

- That is to say we reject $H_0 : \mu_i = \mu_j$ if

$$\left| t^* \right| = \frac{\left| \bar{Y}_i - \bar{Y}_j \right|}{\sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} > t_{(n-k), k(k-1)/2}$$

- If we use $\alpha^* = \frac{\alpha}{k(k-1)/2}$ for each individual test and all the tests are independent then the probability of making at least one error in all $k(k-1)/2$ tests is now approximately $\alpha$.

- However, $\alpha^* = \frac{\alpha}{k(k-1)/2}$ gets very small when $k$ is large.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$k(k-1)/2$</th>
<th>$0.05 / k(k-1)/2$</th>
<th>$1 - \left(1 - \frac{0.05}{k(k-1)/2}\right)^{k(k-1)/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.0500000000</td>
<td>0.05000000</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.0166666667</td>
<td>0.04917130</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0.0083333333</td>
<td>0.04896984</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0.0050000000</td>
<td>0.04889877</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>0.0033333333</td>
<td>0.04885002</td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>0.002380952</td>
<td>0.04882728</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
<td>0.001785714</td>
<td>0.04881309</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>0.001388889</td>
<td>0.04880363</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>0.0011111111</td>
<td>0.04879702</td>
</tr>
</tbody>
</table>

- So, if the tests are not all independent, then this is a bad idea if $k$ is large.

- Why is this bad? When we force the Type I error rate to be very low, we start making Type II errors.

- So, we don’t want to make $\alpha^*$ lower than it needs to be.
Tukey comparisons

- An alternative to Bonferroni comparisons is the Tukey procedure.
- The Tukey procedure ensures that the probability of making at least one error in all \( k(k - 1)/2 \) tests is \( \alpha \), but does not assume that the tests are independent.
- With the Tukey procedure we perform all \( k(k - 1)/2 \) tests, rejecting \( H_0 : \mu_i = \mu_j \) if

\[
|t^*| = \frac{|\bar{Y}_i - \bar{Y}_j|}{\sqrt{MSE \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}} > \frac{q(1 - \alpha, k, n - k)}{\sqrt{2}}
\]

- We find the values of \( q(1 - \alpha, k, n - k) \) in Table A.9 on pp. 914-916 of the text.
The one-way effects model (PNC 9.2)
To this point in the chapter we have used the following model for our data.

\[ Y_{ij} = \mu_i + \epsilon_{ij} \quad \text{for} \quad i = 1, \ldots, k \quad \text{and} \quad j = 1, \ldots, n_i \]

where

\[ \epsilon_{ij} \sim N(0, \sigma^2) \]

The text refers to this as the one-way means model. In Section 9.2, the text introduces an alternative parameterization of this same model called the one-way effects model.

\[ Y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \text{for} \quad i = 1, \ldots, k \quad \text{and} \quad j = 1, \ldots, n_i \]

where

\[ \epsilon_{ij} \sim N(0, \sigma^2) \quad \text{and} \quad \sum_{i=1}^{k} \tau_i = 0 \]

In this model we define \( \mu \) to be an overall mean which we define by

\[ \mu = \frac{1}{k} \sum_{i=1}^{k} \mu_i \]

and \( \tau_i \) is the effect of the \( i \)th population.

\[ \tau_i = \mu_i - \mu \]

In terms of the parameters of this new model, the hypothesis

\[ H_0 : \quad \mu_1 = \mu_2 = \cdots = \mu_k \]
\[ H_a : \quad \text{Not all the } \mu_i \text{ are equal.} \]

can be equivalently stated (and tested) as

\[ H_0 : \quad \tau_1 = \tau_2 = \cdots = \tau_k = 0 \]
\[ H_a : \quad \text{Not all the } \tau_i \text{ are equal to zero.} \]

We estimate \( \mu \) by \( \hat{\mu} = \frac{1}{k} \sum_{i=1}^{k} \bar{Y}_i \) and \( \tau_i \) by \( \hat{\tau}_i = (\bar{Y}_i - \hat{\mu}) \).
The Randomized Complete Block Design

- To this point in our discussion of comparing the means of more than two populations, we have assumed a completely randomized design (CRD). This is an experimental design in which treatments are assigned to experimental units completely at random. In this design, every experimental unit has an equal chance to receive any one of the treatments.

- In MA 2611 (PNC Chapter 3), you also learned about the randomized complete block design (RCBD). This is an experimental design in which the experimental units are divided into blocks and, separately within each block, treatments are assigned at random to the experimental units within that block. In this design, every experimental unit within the same block has an equal chance to receive any one of the treatments.
What is a block?

A block is a group of subjects or experimental units that are similar in ways that are expected to affect the response to the treatments.

Examples of blocks

- Characteristics associated with an experimental unit
  - plants grown in similar areas of the greenhouse
  - fruit flies from the same inbred line
  - patients similar in age, sex, diagnosis, and/or previous treatment
  - corn in neighboring areas of a field

- Characteristics associated with the experimental setting:
  - technician
  - time of processing
  - batch of material
  - measuring instrument

Why block?

- The goal of blocking is to control inherent variation in subject pool by ensuring that each treatment group is somewhat similar.

- Since a block is homogeneous, differences between outcomes are more likely due to the treatments.

- We can more easily look at the effect of the treatments on different homogeneous groups.

- The RCBD forces the treatment groups to be balanced in terms of the extraneous variables.
How does the statistical analysis change for the RCBD?

Statistical model:

\[ Y_{ij} = \mu + \tau_i + \gamma_j + \epsilon_{ij} \quad \text{for} \quad i = 1, \ldots, k \quad \text{and} \quad j = 1, \ldots, b \]

where

\[ \epsilon_{ij} \sim N(0, \sigma^2), \quad \sum_{i=1}^{k} \tau_i = 0 \quad \text{and} \quad \sum_{j=1}^{b} \gamma_j = 0 \]

- \( \mu \) is the overall mean
- \( \tau_i \) is the effect of the \( i \)th treatment
- \( \gamma_j \) is the effect of the \( j \)th block
- \( n = kb \)

Estimation

- The least squares estimators for the parameters of the model are
  \[ \hat{\mu} = \bar{Y}. \]
  \[ \hat{\tau}_i = \bar{Y}_i. - \bar{Y}. \]
  \[ \hat{\gamma}_j = \bar{Y}.j - \bar{Y}. \]

- The fitted values (or predicted values) for the \( Y_{ij} \) are defined
  \[ \hat{Y}_{ij} = \hat{\mu} + \hat{\tau}_i + \hat{\gamma}_j \]
  \[ = \bar{Y}. + (\bar{Y}_i. - \bar{Y}.) + (\bar{Y}.j - \bar{Y}.) \]
  \[ = \bar{Y}_i. + \bar{Y}.j - \bar{Y}. \]

- And the residuals are defined as
  \[ e_{ij} = Y_{ij} - \hat{Y}_{ij} \]
  \[ = Y_{ij} - \bar{Y}_i. - \bar{Y}.j + \bar{Y}. \]
Hypothesis testing

• In terms of the parameters of this model, the hypotheses of interest are

\[ H_0 : \tau_1 = \tau_2 = \cdots = \tau_k = 0 \]
\[ H_{a, \tau} : \text{Not all the population effects } \tau_i \text{ are equal to zero.} \]

and

\[ H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_b = 0 \]
\[ H_{a, \gamma} : \text{Not all the blocking effects } \gamma_j \text{ are equal to zero.} \]

• And we define the test statistics in terms of the ANOVA table.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populations</td>
<td>( k - 1 )</td>
<td>( SSP = b \sum_{i=1}^{k} (Y_{i.} - \bar{Y}.)^2 )</td>
<td>( MSP = SSP/df_P )</td>
</tr>
<tr>
<td>Blocks</td>
<td>( b - 1 )</td>
<td>( SSB = k \sum_{j=1}^{b} (Y_{.j} - \bar{Y}.)^2 )</td>
<td>( MSB = SSB/df_B )</td>
</tr>
<tr>
<td>Error</td>
<td>((k-1)(b-1))</td>
<td>( SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} e_{ij}^2 )</td>
<td>( MSE = SSE/df_E )</td>
</tr>
<tr>
<td>Total</td>
<td>( kb - 1 )</td>
<td>( SSTO = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}.)^2 )</td>
<td></td>
</tr>
</tbody>
</table>

Note:

• \( SSTO = SSP + SSB + SSE \)

• \( df_T = df_P + df_B + df_E \).

We reject \( H_0, \tau \) if

\[ F^{\tau} = \frac{MSP}{MSE} > F_{df_P, df_E, 1-\alpha} \]

We reject \( H_0, \gamma \) if

\[ F^{\gamma} = \frac{MSB}{MSE} > F_{df_B, df_E, 1-\alpha} \]