Hypothesis Testing: Hypotheses

SECTION 4.1

- Hypothesis test
- Null and alternative hypotheses
- Randomization distribution (Section 4.2)

Tea and the Immune System

- L-theanine is an amino acid found in tea
  - Black tea: about 20mg per cup
  - Green tea (standard): varies, as low as 5mg per cup
  - Green tea (shade grown): varies, up to 46mg per cup
  (Shade grown green tea examples: Gyokuro, Matcha)
- Gamma delta T cells are important for helping the immune system fend off infection
- It is thought that L-theanine primes T cells, activating them to a state of readiness and making them better able to respond to future antigens.
- Does drinking tea actually boost your immunity?

Tea and Immune System

The explanatory variable is tea or coffee, and the response variable is immune system response measured in amount of interferon gamma produced. How could we visualize this data?

a) Bar chart
b) Histogram
c) Side-by-side boxplots
d) Scatterplot

Tea and the Immune System

In study comparing tea and coffee and levels of interferon gamma, if tea drinkers have significantly higher levels of interferon gamma, can we conclude that drinking tea rather than coffee caused an increase in this aspect of the immune response?

a) Yes
b) No
The explanatory variable is tea or coffee, and the response variable is immune system response measured in amount of interferon gamma produced. How might we summarize this data?

- Mean
- Proportion
- Difference in means
- Difference in proportions
- Correlation

A hypothesis test uses data from a sample to assess a claim about a population.

Hypotheses

- Null Hypothesis ($H_0$): Claim that there is no effect or difference.
- Alternative Hypothesis ($H_a$): Claim for which we seek evidence.

Hypothesis Test

- One mean is higher than the other *in the sample*
- Is this difference large enough to conclude the difference is real, and holds for the true population parameters?

Tea and Immune System

- Null Hypothesis ($H_0$): No difference between drinking tea and coffee regarding interferon gamma
  - No "effect" or no "difference"
- Alternative Hypothesis ($H_a$): Drinking tea increases interferon gamma production more than drinking coffee
  - Claim we seek "evidence" for
Hypotheses: parameters

• More formal hypotheses:
  - $\mu_T =$ true mean interferon gamma response after drinking tea
  - $\mu_C =$ true mean interferon gamma response after drinking coffee

$$H_0: \mu_T = \mu_C$$
$$H_a: \mu_T > \mu_C$$

Hypotheses: parameters

• More formal hypotheses:
  - $\mu_T =$ true mean interferon gamma response after drinking tea
  - $\mu_C =$ true mean interferon gamma response after drinking coffee

$$H_0: \mu_T = \mu_C$$
$$H_a: \mu_T > \mu_C$$

Difference in Hypotheses

• Note: the following two sets of hypotheses are equivalent, and can be used interchangeably:

$$H_0: \mu_1 = \mu_2$$
$$H_a: \mu_1 \neq \mu_2$$

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_a: \mu_1 - \mu_2 \neq 0$$

Alternative Hypothesis

If the researchers were simply comparing tea and coffee, with no a priori hypothesis about which would yield a higher immune response, what would the alternative hypothesis be?

a) $H_a: \mu_T = \mu_C$

b) $H_a: \mu_T < \mu_C$

c) $H_a: \mu_T > \mu_C$

d) $H_a: \mu_T \neq \mu_C$

Hypothesis Helpful Hints

• Hypotheses are always about population parameters, not sample statistics

• The null hypothesis always contains an equality

• The alternative hypothesis always contains an inequality ($<, >, \neq$)

• The type of inequality in the alternative comes from the wording of the question of interest

Statistical Hypotheses

Usually the null is a very specific statement

Can we reject the null hypothesis?

Two Plausible Explanations

• If the sample data support the alternative, there are two plausible explanations:
  1. The alternative hypothesis ($H_a$) is true
  2. The null hypothesis ($H_0$) is true, and the sample results were just due to random chance

• Key question: Do the data provide enough evidence to rule out #2?
Two Plausible Explanations

- Why might the tea drinkers have higher levels of interferon gamma?
- Two plausible explanations:
  - **Alternative true**: Tea causes increase in interferon gamma production
  - **Null true, random chance**: the people who got randomly assigned to the tea group have better immune systems than those who got randomly assigned to the coffee group

The Plausibility of the Null

- The goal is determine whether the null hypothesis and random chance are a plausible explanation, given the observed data
- Key idea: How unlikely would it be to see a sample statistic as extreme as we’ve observed, just by random chance, if the null hypothesis were true?
- How do we figure this out?
  - **SIMULATE what would happen if H₀ were true!**

Tea and Immune Response

1. Randomize units to treatment groups
2. Conduct experiment
3. Measure response variable
4. Calculate statistic

Tea and Immune Response

<table>
<thead>
<tr>
<th>Tea</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>48</td>
</tr>
<tr>
<td>18</td>
<td>51</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>11</td>
<td>46</td>
</tr>
<tr>
<td>18</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>31</td>
<td>38</td>
</tr>
<tr>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>56</td>
<td>52</td>
</tr>
<tr>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>58</td>
<td>52</td>
</tr>
</tbody>
</table>

- **x_T** = 17.12
- **x_C** = 17.12
Measuring Evidence against $H_0$

To see if a statistic provides evidence against $H_0$, we need to see what kind of sample statistics we would observe, just by random chance, if $H_0$ were true.

Simulation

• "by random chance" means the random assignment to the two treatment groups
• "if $H_0$ were true" means that interferon gamma levels would be the same, regardless of whether you drink tea or coffee
• To simulate what would happen just by random chance, if $H_0$ were true...
• Re-randomize units to treatment groups, keeping the response values unchanged

Tea and Immune Response

<table>
<thead>
<tr>
<th>Tea</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td>42</td>
<td>27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tea</th>
<th>Coffee</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>55</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>48</td>
</tr>
<tr>
<td>12</td>
<td>52</td>
</tr>
</tbody>
</table>

Simulation

Repeat Many Times!

1. Re-randomize units to treatment groups
2. Calculate statistic:

$\bar{x}_T - \bar{x}_C = 12.3$

Distribution of Statistic Under $H_0$

How extreme is the observed statistic???

Is the null hypothesis a plausible explanation?
Randomization Distribution

A randomization distribution is a collection of statistics from samples simulated assuming the null hypothesis is true.

- The randomization distribution shows what types of statistics would be observed, just by random chance, if the null hypothesis were true.

Green Tea and Prostate Cancer

- A study was conducted on 60 men with PIN lesions, some of which turn into prostate cancer.
- Half of these men were randomized to take 600 mg of green tea extract daily, while the other half were given a placebo pill.
- The study was double-blind; neither the participants nor the doctors knew who was actually receiving green tea.
- After one year, only 1 person taking green tea had gotten cancer, while 9 taking the placebo had gotten cancer.

In the study about green tea and prostate cancer, if the difference is statistically significant, could we conclude that green tea really does help prevent prostate cancer?

(a) Yes
(b) No

The explanatory variable is green tea extract of placebo, the response variable is whether or not the person developed prostate cancer. What statistic and parameter is most relevant?

- Mean
- Proportion
- Difference in means
- Difference in proportions
- Correlation

$p_1 =$ proportion of green tea consumers to get prostate cancer  
$p_2 =$ proportion of placebo consumers to get prostate cancer

State the null hypotheses.

- $H_0: p_1 = p_2$
- $H_0: p_1 < p_2$
- $H_0: p_1 > p_2$
- $H_0: p_1 \neq p_2$

$p_1 =$ proportion of green tea consumers to get prostate cancer  
$p_2 =$ proportion of placebo consumers to get prostate cancer

State the alternative hypotheses.

- $H_a: p_1 = p_2$
- $H_a: p_1 < p_2$
- $H_a: p_1 > p_2$
- $H_a: p_1 \neq p_2$
Randomization Test

1. State hypotheses
2. Collect data
3. Calculate statistic: \( \hat{p}_1 - \hat{p}_2 = 0.033 - 0.300 = 0.267 \)
4. Simulate statistics that could be observed, just by random chance, if the null hypothesis were true (create a randomization distribution)
5. How extreme is the observed statistic?
6. Is the null hypothesis (random chance) a plausible explanation?

Randomization Distribution

Based on the randomization distribution, would the observed statistic be extreme if the null hypothesis were true?

- a) Yes
- b) No

Randomization Distribution

Do you think the null hypothesis is a plausible explanation for these results?

- a) Yes
- b) No

Randomization Distribution Center

A randomization distribution simulates samples assuming the null hypothesis is true, so

- A randomization distribution is centered at the value of the parameter given in the null hypothesis.

In a hypothesis test for \( H_0: \mu = 12 \) vs \( H_a: \mu < 12 \), we have a sample with \( n = 45 \) and \( \bar{x} = 10.2 \).

What do we require about the method to produce randomization samples?

- a) \( \mu = 12 \)
- b) \( \mu < 12 \)
- c) \( \bar{x} = 10.2 \)
In a hypothesis test for $H_0: \mu = 12$ vs $H_a: \mu < 12$, we have a sample with $n = 45$.

What will we look for on the randomization distribution?

a) How extreme 10.2 is
b) How extreme 12 is
c) How extreme 45 is
d) What the standard error is
e) How many randomization samples we collected

In a hypothesis test for $H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 > \mu_2$, we have a sample with 26 and 21.

What do we require about the method to produce randomization samples?

a) $\mu_1 = \mu_2$
b) $\mu_1 > \mu_2$
c) $\bar{x}_1 = 26, \bar{x}_2 = 21$
d) $\bar{x}_1 - \bar{x}_2 = 5$

In a hypothesis test for $H_0: \mu_1 = \mu_2$ vs $H_a: \mu_1 > \mu_2$, we have a sample with 26 and 21.

Where will the randomization distribution be centered?

a) 0
b) 1
c) 21
d) 26
e) 5

In a hypothesis test for $H_a: \mu_1 = \mu_2$ vs $H_a: \mu_1 > \mu_2$, we have a sample with 26 and 21.

What do we look for on the randomization distribution?

a) The standard error
b) The center point
c) How extreme 26 is
d) How extreme 21 is
e) How extreme 5 is

Summary

- Hypothesis tests use data from a sample to assess a claim about a population.
- Hypothesis tests are usually formalized with competing hypotheses:
  - Null hypothesis ($H_0$): no effect or no difference
  - Alternative hypothesis ($H_a$): what we seek evidence for
- We assess whether the null hypothesis is plausible by:
  1. seeing what kinds of statistics we would observe by random chance, if the null hypothesis were true
  2. assessing the extremity of our observed statistic

To Do

- Read Section 4.1
- HW 4.1 due Friday, 10/16
Null Hypothesis

http://xkcd.com/892/