The Scientific Method and Hypothesis Testing

For the record...

The term data is plural.

“Collecting these data required considerable personal expenditure.”

The scientific method is used because it is defensible AND repeatable.

- The scientific method is a defensible technique that helps to minimize external influences.
- Statistical analyses are repeatable.
- The scientific method consists of hypotheses, models, laws, and theories.

Hypothesis – a proposition or statement whose truth or falsity is capable of being tested.

Models – a means of simplifying reality so that relationships between variables can be more clearly studied.

Laws – universal statements of unrestricted range.

Theories – a collection of laws that gives greater insight than produced by individual laws.
Statistical procedures for addressing research questions involves formulating a concise statement of the hypothesis to be tested.

The hypothesis to be tested is referred to as the null hypothesis (abbreviated $H_0$) because it is a statement of no difference.

Hypothesis testing starts with the assumption that the null hypothesis is true... that there is/are no difference(s).

Along with the null hypothesis we must also state an alternate hypothesis (abbreviated $H_a$).

The alternate hypothesis is a statement that a difference exists.

If a null hypothesis is rejected, then we tentatively accept the alternate hypothesis and conclude that there is a difference.

Why is the null hypothesis the one that is tested?

Think about it this way: we only have to find one instance in which the null hypothesis is not true (false) in order to be able to reject it.

Conversely, we would have to continue to test the alternate hypothesis in order to be able to accept it.

In other words we would have to test all possibilities since the alternate hypothesis can only be proven correct if all possible tests are performed.

The moral of the story:

It is easier to prove a null hypothesis incorrect than to prove an alternate hypothesis correct.
Example Hypotheses:

$H_0$: There is no difference in eye color between the two groups.

$H_a$: There is a difference in eye color between the two groups.

IMPORTANT:

In ALL cases, if the calculated value is GREATER than the critical value (from the table) then reject $H_0$ and accept $H_a$.

If the calculated value is LESS than the critical value then accept $H_0$.

When a statistical test is performed there are two distinct, but related results:

• Test statistic: this is the value that is calculated.

• Probability: this is the probability of having that particular test statistic given the characteristics of the data set.

These characteristics are values such as the mean, standard deviation, sample size, etc...

Which characteristics are important depends on the specific test.

The results of any statistical test (e.g. one where you are testing a null hypothesis) must be stated in a concise summary statement. This statement should include a summary of the findings, the test that was performed, the alpha level used, the statistical results, and the probability.
Alpha Level ($\alpha$)

The alpha level is the probability of committing a Type I error.
- A true null hypothesis can be incorrectly rejected.
- Also called alpha or significance level.
- It is essentially a 'false positive'.

By convention we typically use 0.05 or 0.01 (5% or 1%) as our alpha level.

Beta Level ($\beta$)

The beta level is the probability of committing a Type II error.
- A false null hypothesis we fail to reject.
- It is essentially a 'false negative'.

By convention this value is not specified.

Type 1 and 2 Errors

<table>
<thead>
<tr>
<th>If $H_0$ is true</th>
<th>If $H_a$ is false</th>
</tr>
</thead>
<tbody>
<tr>
<td>And $H_0$ is rejected:</td>
<td>Type I Error</td>
</tr>
<tr>
<td>And $H_0$ is not rejected:</td>
<td>No Error</td>
</tr>
</tbody>
</table>

A Type I error accepts an alternate hypothesis when the results can be attributed to chance.
- So in effect we are stating that there is a difference when none actually exists.

A Type II error is only an error in the sense that we fail to correctly reject a false null hypothesis.
- It is not an error in the sense that an incorrect conclusion was drawn since no conclusion is drawn when the null hypothesis is accepted. It is the lesser of two evils.

Alpha level of 0.05

An $\alpha$ of 0.05 is equal to ±1.96 sd from the mean.
One and Two-tailed Tests

Two-tailed statistics test for difference but not direction.

One-tailed statistics test for difference and direction.

Earlier research suggested that the average house value in York was lower than that in Lancaster.

Data for the average house value for each block group in downtown York and Lancaster were gathered from the census.

A two-sample t-test was performed to determine whether housing values were lower in York than in Lancaster.

Since we have a priori (prior) knowledge of the direction of the difference (e.g. housing values are less in York) we would use a one-tailed test.
H$_{0}$: Housing values in York are not significantly less than those in Lancaster.

H$_{a}$: Housing values in York are significantly less than those in Lancaster.

Note how the direction of the difference is stated in the hypotheses.

The test we will use is called a T-Test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{s_1^2 + s_2^2}$$

$$s_1^2 = \frac{SS_1}{v_1}$$

$$s_2^2 = \frac{SS_2}{v_2}$$

$$SS_1 = \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2$$

$$SS_2 = \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

$$\alpha = 0.05$$

$$n_1 = 7$$  $$n_2 = 6$$

df = (n_1 + n_2 - 2) = (7 + 6 - 2) = 11

$$v_1 = 7 - 1 = 6$$

$$v_2 = 6 - 1 = 5$$

Housing Value (s)

<table>
<thead>
<tr>
<th>York</th>
<th>Lancaster</th>
</tr>
</thead>
<tbody>
<tr>
<td>25368</td>
<td>49465</td>
</tr>
<tr>
<td>37045</td>
<td>37500</td>
</tr>
<tr>
<td>47500</td>
<td>53055</td>
</tr>
<tr>
<td>26785</td>
<td>48125</td>
</tr>
<tr>
<td>41493</td>
<td>45000</td>
</tr>
<tr>
<td>32864</td>
<td>52946</td>
</tr>
<tr>
<td>26140</td>
<td></td>
</tr>
</tbody>
</table>

$$\alpha = 0.05$$

$$n_1 = 7$$  $$n_2 = 6$$

df = (n_1 + n_2 - 2) = (7 + 6 - 2) = 11

$$v_1 = 7 - 1 = 6$$

$$v_2 = 6 - 1 = 5$$

$$\bar{X}_{York} = \frac{25368 + 37045 + 47500 + 26785 + 41493 + 32864 + 26140}{7} = \frac{212202}{7} = 3031.43$$

$$\bar{X}_{Lancaster} = \frac{49465 + 37500 + 53055 + 48125 + 45000 + 52946}{6} = \frac{282330}{6} = 4705.50$$

$$SS_{York} = (25368 - 3031.43)^2 + (37045 - 3031.43)^2 + (47500 - 3031.43)^2 + (26785 - 3031.43)^2 + (41493 - 3031.43)^2 + (32864 - 3031.43)^2 + (26140 - 3031.43)^2 = 47212244$$

$$SS_{Lancaster} = (49465 - 4705.50)^2 + (37500 - 4705.50)^2 + (53055 - 4705.50)^2 + (48125 - 4705.50)^2 + (45000 - 4705.50)^2 + (52946 - 4705.50)^2 + (4705.50 - 4705.50)^2 = 47212244$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{SS_{York} + SS_{Lancaster}}{v_1 + v_2}} = \sqrt{\frac{47212244 + 47212244}{11}} = \sqrt{161255288} = 12712.46$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{3031.43 - 4705.50}{12712.46} = -1.532$$

$$t_{critical} = 1.796$$

Since 1.532 < 1.796 accept $H_0$.
The calculated $t$ value ($1.532$) is then compared to the critical $t$ value ($1.796$).

- If it is higher then we **REJECT $H_0$**.
- If it is lower then we **ACCEPT $H_0$**.

The critical values are either taken from a table or calculated in SPSS.

Our **HIGH POWER** summary statement would then be:

The housing values in York were not significantly less than the housing values in Lancaster ($t_{0.05,1.532} < 1.00 > p > 0.05$).

---

Every time a statistical test is performed, you need to include the following:

1. **A statement of the alpha level**.
2. **Null and alternate hypotheses**.
3. **A high power summary statement that contains**:
   - The test performed.
   - The calculated test statistic.
   - The exact probability or probability range.

**Developing Hypotheses Based on Maps**

Maps can be used for:
- Delineating groups
- Locating observations
- Determining measurements

Measurements can be (but are not limited to):
- Elevation
- Azimuth
- Aspect
- Distance
- Proximity
Develop null and alternate hypotheses based the following maps.
Often our hypotheses are not concerned with differences between or among groups.

- For example, an association between a measured variable and a measured landscape or natural characteristic.
- e.g. Decreasing temperature with increasing altitude.

In such cases, our null and alternate hypotheses might be:

\[ H_0: \text{There no inverse association between temperature and altitude.} \]
\[ H_a: \text{There is an inverse association between temperature and altitude.} \]

or more simply:

\[ H_0: \text{There no association between temperature and altitude.} \]