A Trial Analogy for Statistical Hypothesis Testing

<table>
<thead>
<tr>
<th>Legal Trial</th>
<th>Statistical Significance Test</th>
</tr>
</thead>
</table>
| • Begin with claim:  
  – Smith is not guilty  
  If this is rejected, we accept  
  – Smith is guilty  
  • “reasonable doubt”  
  • Present evidence (facts)  
  • Evaluate evidence | • Hypotheses (statements)  
  – Ho: Null hypothesis  
  – $H_1$: Alternative hypothesis  
  • Alpha level  
  • Collect data  
  • Statistical evaluation according to prescribed rules (test statistic) |
A Trial Analogy for Statistical Hypothesis Testing

<table>
<thead>
<tr>
<th>Trial</th>
<th>Hypothesis Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verdict</strong></td>
<td><strong>Statistical Decision</strong></td>
</tr>
<tr>
<td>– Guilty beyond reasonable doubt</td>
<td>– Beyond alpha-level, reject Ho.</td>
</tr>
<tr>
<td>– Not guilty beyond reasonable doubt</td>
<td>– Within alpha-level, fail to reject Ho.</td>
</tr>
<tr>
<td><strong>Legal Mistakes</strong></td>
<td><strong>Types of Errors</strong></td>
</tr>
<tr>
<td>– Execute Smith when he is innocent</td>
<td>– Type I: reject Ho, when Ho is TRUE</td>
</tr>
<tr>
<td>– Free Smith when he is guilty</td>
<td>– Type II: fail to eject Ho when it is FALSE.</td>
</tr>
</tbody>
</table>
# Types of Errors

<table>
<thead>
<tr>
<th>Decision</th>
<th>True State of Nature</th>
<th>Type I error (rejecting a true null hypothesis) $\alpha$</th>
<th>Type II error (failing to reject a false null hypothesis) $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>We decide to reject the null hypothesis.</td>
<td>The null hypothesis is true.</td>
<td>Correct decision</td>
<td></td>
</tr>
<tr>
<td>We fail to reject the null hypothesis.</td>
<td>The null hypothesis is false.</td>
<td>Correct decision</td>
<td></td>
</tr>
</tbody>
</table>

| Table 7-1 | Type I and Type II Errors |
Example: Assume that we are conducting a hypothesis test of the claim $p > 0.5$. Here are the null and alternative hypotheses: $H_0: p = 0.5$, and $H_1: p > 0.5$.

a) Identify a type I error.

Conclude that there is sufficient evidence to support $p > 0.5$, when in reality $p = 0.5$.

b) Identify a type II error.

Fail to reject $p = 0.5$ (and therefore fail to support $p > 0.5$) when in reality $p > 0.5$. 
Null Hypothesis: $H_0$

- The \textit{null hypothesis} includes the assumed value of the population parameter.
- It must be a statement of equality.
- Test the Null Hypothesis \textbf{directly}
- \textbf{Reject} $H_0$ or fail to reject $H_0$
Alternative Hypothesis: $H_a$

- The *alternative hypothesis* (denoted by $H_1$ or $H_a$) is the statement that the parameter has a value that somehow differs from the null hypothesis.

- $\neq, <, >$
Example: Identify the Null and Alternative Hypothesis. Refer to Figure 7-2 and use the given claims to express the corresponding null and alternative hypotheses in symbolic form.

a) The proportion of drivers who admit to running red lights is greater than 0.5.

b) The mean height of professional basketball players is at most 7 ft.

c) The standard deviation of IQ scores of actors is equal to 15.
The test statistic is a value computed from the sample data, and it is used in making the decision about the rejection of the null hypothesis.

\[ z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \]

Test statistic for proportion=p, q=1-p
Example: A survey of $n = 880$ randomly selected adult drivers showed that $56\%$ (or $\hat{p} = 0.56$) of those respondents admitted to running red lights. Find the value of the test statistic for the claim that the majority of all adult drivers admit to running red lights. (In Section 7-3 we will see that there are assumptions that must be verified. For this example, assume that the required assumptions are satisfied and focus on finding the indicated test statistic.)
Solution: The preceding example showed that the given claim results in the following null and alternative hypotheses: $H_0: \ p = 0.5$ and $H_1: \ p > 0.5$. Because we work under the assumption that the null hypothesis is true with $p = 0.5$, we get the following test statistic:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{0.56 - 0.5}{\sqrt{(0.5)(0.5)/880}} = 3.56$$
Unusually high sample proportions

Critical region:
Area of $\alpha = 0.05$ used as criterion for identifying unusually high sample proportions

$p = 0.5$

$z = 1.645$

Critical value

Sample proportion of: $\hat{\rho} = 0.56$

Test Statistic $z = 3.56$

Proportion of adult drivers admitting that they run red lights

P-value = probability that the test Statistic would be more extreme than the observed value

$= P(Z > 3.56) < 1 - 0.999767 = 0.000233$
The $P$-value (or $p$-value or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true. The null hypothesis is rejected if the $P$-value is very small, such as 0.05 or less.
Conclusions in Hypothesis Testing

We always test the null hypothesis.

1. Reject the $H_0$

2. Fail to reject the $H_0$
Two-tailed Test

$H_0: = \neq$ 

$H_a: \neq$ 

$\alpha$ is divided equally between the two tails of the critical region

Means less than or greater than

Reject $H_0$

Fail to reject $H_0$

Reject $H_0$

Sign used in $H_1: \neq$
Right-tailed Test

$H_0: =$

$H_a: >$

Points Right

Fail to reject $H_0$

Reject $H_0$

Sign used in $H_1: >$
Left-tailed Test

\[ H_0: = \]
\[ H_a: < \]

Points Left

Reject \( H_0 \)  
Fail to reject \( H_0 \)

Sign used in \( H_1: < \)
Example: Finding P-values.

Start

What type of test?

Two-tailed

Is the test statistic to the right or left of center?

Left-tailed

P-value = area to the left of the test statistic

P-value = twice the area to the left of the test statistic

P-value = twice the area to the right of the test statistic

P-value = area to the right of the test statistic
Example: Finding $P$-values. First determine whether the given conditions result in a right-tailed test, a left-tailed test, or a two-tailed test, then find the $P$-values and state a conclusion about the null hypothesis.

a) A significance level of $\alpha = 0.05$ is used in testing the claim that $p > 0.25$, and the sample data result in a test statistic of $z = 1.18$.

b) A significance level of $\alpha = 0.05$ is used in testing the claim that $p \neq 0.25$, and the sample data result in a test statistic of $z = 2.34$. 
Traditional method:

*Reject $H_0$* if the test statistic falls within the critical region.

*Fail to reject $H_0$* if the test statistic does not fall within the critical region.
Decision Criterion

$P$-value method:

*Reject $H_0$ if $P$-value $\leq \alpha$ (where $\alpha$ is the significance level, such as 0.05).*

Fail to *reject* $H_0$ if $P$-value $> \alpha$. 
Another option:

Instead of using a significance level such as 0.05, simply identify the $P$-value and leave the decision to the reader.
Decision Criterion

Confidence Intervals:

Because a confidence interval estimate of a population parameter contains the likely values of that parameter, reject a claim that the population parameter has a value that is not included in the confidence interval.
Wording of Final Conclusion

Start

- Does the original claim contain the condition of equality?
  - Yes (Original claim contains equality)
    - Do you reject $H_0$?
      - Yes (Reject $H_0$)
        - Wording of final conclusion: "There is sufficient evidence to warrant rejection of the claim that . . . (original claim)."
        - (This is the only case in which the original claim is rejected.)
      - No (Fail to reject $H_0$)
        - Wording of final conclusion: "There is not sufficient evidence to warrant rejection of the claim that . . . (original claim)."
        - (This is the only case in which the original claim is supported.)
  - No (Original claim does not contain equality and becomes $H_1$)
    - Do you reject $H_0$?
      - Yes (Reject $H_0$)
        - Wording of final conclusion: "The sample data support the claim that . . . (original claim)."
      - No (Fail to reject $H_0$)
        - Wording of final conclusion: "There is not sufficient sample evidence to support the claim that . . . (original claim)."