Chapter 1. Making decisions

Population, Sample

Statistical inference

Hypothesis's

- Null hypothesis $H_0$
- Alternative hypothesis $H_1$

Note: $H_0$, $H_1$ are statements about the same population.

Statistically significant (supporting $H_1$)

$\rightarrow$ Reject $H_0$

1.3.3 What errors could we make?

Definition:

- **Type I error**: rejecting $H_0$ when in fact it is true.
- **Type II error**: accepting $H_0$ when in fact it is not true.

<table>
<thead>
<tr>
<th></th>
<th>$H_0$ is true</th>
<th>$H_1$ is true</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>accept $H_0$</strong></td>
<td>no error</td>
<td>Type II</td>
</tr>
<tr>
<td><strong>reject $H_0$</strong></td>
<td>Type I</td>
<td>no error</td>
</tr>
</tbody>
</table>
Tip: Type I error may only happen when \( H_0 \) is true.

How to select \( H_0 \)?
- \( H_0 \): \text{innocent}.
- \( H_1 \): \text{guilty}.

No type I error \( \rightarrow \) never reject \( H_0 \).

Ex. 1.4 - Rain
- \( H_0 \): Tonight it is going to rain.
- \( H_1 \): Tonight it is not going to rain.

Type I error: In fact it is going to rain, but you decide that it is not.
   - didn't bring umbrella.
   - got wet.

Type II error: In fact it is not going to rain but you decided that it is.
   - did bring umbrella.
Let's do it. 1.5 Testing a new drug

Ho: The new drug is as effective.
Hi: The new drug is more effective.

Type I error:
Conclusion that the new drug is better when it's not.

Type II error:
Concluded that the new drug is no better when it is actually better.

\[ \alpha = \text{level of significance} \]
\[ = \text{the chance of a Type I error occurring} \]
\[ = \text{the chance of rejecting } H_0 \text{ when it's true} \]
\[ \beta = \text{the chance of a Type II error occurring} \]
\[ = \text{the chance of accepting } H_0 \text{ when it is not true} \ (H_1 \text{ is true}) \]
Type II error = accept $H_0$ when $H_1$ is true

= decide that the shown bag is Bag A when it is Bag B

= (kept the other bag, Bag A)
  must pay $560

<table>
<thead>
<tr>
<th>Free Value</th>
<th>Bag A</th>
<th>Bag B</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1000</td>
<td>1/20</td>
<td>0</td>
</tr>
<tr>
<td>$10</td>
<td>7/20</td>
<td>1/20</td>
</tr>
<tr>
<td>$20</td>
<td>6/20</td>
<td>1/20</td>
</tr>
<tr>
<td>$50</td>
<td>2/20</td>
<td>2/20</td>
</tr>
<tr>
<td>$40</td>
<td>2/20</td>
<td>2/20</td>
</tr>
<tr>
<td>$50</td>
<td>1/20</td>
<td>6/20</td>
</tr>
<tr>
<td>$60</td>
<td>1/20</td>
<td>7/20</td>
</tr>
<tr>
<td>$1000</td>
<td>0</td>
<td>1/20</td>
</tr>
</tbody>
</table>

**Definition:**

The **direction of extreme** corresponds to the position of the values that are more likely under the alternative hypothesis $H_1$ than under the null hypothesis $H_0$.

If larger values are more likely under $H_1$ than the direction of extreme is said to be to the **right**.
Decision Rule 1:
Reject \( H_0 \) if you select a $60 or $1000 voucher, otherwise accept \( H_0 \).
\[
\begin{align*}
& \geq 60 \quad \text{reject } H_0 \\
& < 60 \quad \text{accept } H_0.
\end{align*}
\]

\( \alpha \) = significance level
= chance of rejecting \( H_0 \) when \( H_0 \) is true

= chance of selecting $60 or $1000 from Bag A.
= \( \frac{1}{20} = 0.05 \)

\( \beta \) = chance of accepting \( H_0 \) when \( H_1 \) is true

= chance of selecting < $60 from Bag B.
= \( \frac{12}{20} = 0.6 \)
Definition:
A rejection region is the set of values for which you would reject $H_0$. Such values are favor the $H_1$.

An acceptance region is the set of values for which you would accept $H_0$.

The cut-off value or critical value is the value which marks the starting point of the set of values that comprise the rejection region.

Decision rule #1: $60$

cut-off value should work with direction of extreme

Decision Rule 2:
Reject $H_0$ if $x \geq 60$. $< 60$ accept $H_0$.

$x$ = chance of selecting $\geq 60$ from bag A.
$= \frac{3}{20} = 0.1$

$\beta$ = chance of selecting $< 60$ from bag B.
$= \frac{17}{20} = 0.3$
Decision Rule 3:

Reject $H_0$ if $\bar{y} \geq \$40$

$\alpha = 0.05$ accept $H_0$

$\beta = < \$40$ from Bag A

$= \frac{4}{20} = 0.2$

$\beta = < \$40$ from Bag B

$= \frac{4}{20} = 0.2$

Summary of the relationship between
the decision rule and the significance level

Decision Rule $\rightarrow$ Significance Level
(given cutoff value)

Significance Level $\rightarrow$ Decision Rule
(find cutoff value)

$\alpha = 0.1$ given $\rightarrow$ decision rule 2

Definition:
A rejection region is called one-sided if
the set of extreme values are all in one direction.
A rejection region is called two-sided if
the set of extreme values are in two directions,
both to the right and to the left.