New Topic: Hypothesis Testing

Hypothesis: Fact either true or false.

E.g.: This coin is fair. I.e. $X \sim \text{Bernoulli}$ with $\theta$. Fair is if $\theta = 1/2$.

Innocent until proven guilty.

Null hypothesis: $H_0$: "nothing to see here".

"Coin is fair" $\iff$ this drug does not improve outcomes.

Test $H_0$ against alternative hypothesis $H_1$: not necessarily $\sim H_0$.

E.g.: Pharma manufacturers. Proportion of population that gets a disease when exposed to a virus is $90\%$. Test hypothesis that their drug decreases proportion of diseased to $60\%$.

$\Phi$: Population. People exposed to the virus & given drug.

$H_0$: the drug does not help. $X \sim \text{Bernoulli}$ with $\theta = 0.9$.
$H_1$: alternative: $X_\theta = \text{Bernoulli with } \theta > 0.6$.

Tests give $n=20$ people dry. Reject $H_0$ if $k = \# \text{ get disease is } k \leq 14$. Q: how good is this test?

**Def:** Critical region/value: set of outcomes for which we reject $H_0$.

```
<table>
<thead>
<tr>
<th>Ho true</th>
<th>Ho false</th>
</tr>
</thead>
<tbody>
<tr>
<td>✅</td>
<td>✅</td>
</tr>
</tbody>
</table>
```

**Type I error:** accept $H_0$ but $H_1$ is true.

$\Pr(\text{Type I error}) = \alpha$, $\Pr(\text{Type II error}) = \beta$.

\[ x = \Pr(\text{Type I error}) = \Pr(H_0 \text{ is true } \& \text{ reject } H_0) \]
\[ = \Pr(\theta = 0.9 \& n\bar{X} \leq 14) = \sum_{k=0}^{14} \binom{20}{k} 0.9^k (1-0.9)^{20-k} = 0.011. \]

$\alpha = 1.1\%$: "size of critical region" $\Rightarrow$ "level of significance".

\[ \beta = \Pr(\text{Type II error}) = \Pr(H_1 \text{ is true } \& \text{ accept } H_0) \]
\[ = \Pr(\theta = 0.6 \& n\bar{X} > 15) = \sum_{k=15}^{20} \binom{20}{k} 0.6^k (1-0.6)^{20-k} = 0.1256. \]

$\Rightarrow 12\%$. 
Brachistochrone: What is the path of shortest time for a ball to roll from A to B.

(varational analysis). (only gravity).

\[ \Phi \]

\[ \begin{array}{c}
-1 \\
0 \\
1 \\
\end{array} \]  \[ \begin{array}{c}
m \\
\sigma \\
\end{array} \]

E.g.: normal population \( X_1, \ldots, X_n \), \( \text{var} \sigma^2 = 1 \). \( \bar{X} \)

Test \( H_0: M = M_0 \) against \( H_1: M = M_1 \), \( \leq \) two simple hypotheses.

When the hypothesis completely determines underlying pdf, that hypothesis is called sample. Otherwise hypothesis is composite, e.g. \( \theta < 0.4 \).

Q: Find a value of \( K \) s.t. \( \bar{X} > K \) is critical region has level of significance \( \alpha = 0.05 \).

\[ \alpha = P(\text{Type I error}) = P(\text{Ho rejected} \ & \ \text{true}) \]

\[ = P(M = M_0 \ & \ \bar{X} > K) \]

\( \bar{X} \sim N(M_0, \sigma^2 = \frac{1}{n}) \).
\[ P(F \geq K) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi n}} e^{-\frac{(z-m_0)^2}{2/n}} \, dz \]

Need to convert to standard normal so we can "look up" \( Z \chi \):

\[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \, dz = 1 \]

Let \( u^2 = \frac{(z-m_0)^2}{2/n} \), then \( u = \frac{(z-m_0)\sqrt{n}}{\sqrt{2}} \).

\[ \frac{dz}{\sqrt{n}} = \frac{1}{\sqrt{2}} \, du \]

\[ -\frac{u^2}{2} \]

\[ \int \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du = \phi \]

\[ (K-m_0)\sqrt{n} = \phi \chi \]

\[ K = m_0 + \frac{\phi \chi}{\sqrt{n}} \]

---

**Part 2:** Is \( H_1 \) "good" i.e. what is \( \beta = P(\text{Type II error}) \)?

Def.: When testing \( H_0 \), the alternative \( H_1 \) has **power** \( 1-\beta \).

I.e. Find power of \( H_1 \).
\[ \beta = P( \text{Type II error} ) = P( \mu = \mu_1, \bar{x} < K ). \]
\[ = P( \frac{\bar{x} - \mu_1}{\sqrt{\frac{\sigma}{n}}} < \frac{(K - \mu_1)}{\sqrt{\frac{\sigma}{n}}} ). \]
\[ = P( Z < \frac{2\sigma}{\sigma_0 + \sqrt{2\sigma}} - \mu_1). \]

**Recap:** K was chosen to give a critical region of level of significance \( \alpha \), so \( K = (\alpha, \mu_0) \). Then the power of the test is \( 1 - \beta \), where \( \beta = \int_{-\infty}^{(\mu_0 - \mu_1)\frac{\sigma}{\sigma_0 + \sqrt{2\sigma}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du \).

**Q:** How many \( n \) to test if we want \( \beta \leq 0.06 \), for \( \mu_0 = 10, \mu_1 = 11 \) (\( \alpha = 0.05 \)).

\[ (\mu_0 - \mu_1)\frac{\sigma}{\sigma_0 + \sqrt{2\sigma}} \leq -1.555 \]
\[ -\frac{\sigma}{\sigma_0 \sqrt{2\sigma}} + 1.645 \leq -1.555 \]
\[ 3.2 \leq \frac{\sigma}{\sigma_0 + \sqrt{2\sigma}} \leq 11 \]
\[ 0.06 \]
\[ -2 \cdot 0.06 = -1.555 \]

\[ n \geq 11 \]