Hypothesis Testing

• Chapter 8

• Chapter 9 - section 10

• Testing hypotheses with data

• The case of Madam X
  • professes an ability to predict the future
  • devise a means of testing her ability and deciding whether her claims are valid
Hypothesis Testing

• Logic of hypothesis testing

• Assume some hypothesis is true

  Madam X has no special power

• Establish alternative hypothesis

  Madam X can predict better than chance

• Collect a sample of data and determine whether the observed result is unlikely under the assumed hypothesis (null hypothesis; $H_0$)

  • if result is unlikely, then reject the null hypothesis and accept the alternative hypothesis ($H_1$)

  • if result is not unlikely, then …
Hypothesis Testing

- Null hypothesis
- Coin flip predictions – series of 12 coin flips
- Distribution of outcomes (number of correct predictions) for a person who is just guessing

\[ p(6) = (924)(.5)^{12} \]

\[ = .2256 \]

\[ p(6) = .0002 + .0029 + .0161 = .0192 \]
Hypothesis Testing

• Null hypothesis significance testing
  • provides the probability of the observed outcome (or one that is more extreme): \( p(\text{Data} \mid H_0) \)

• coin flip predictions: \( p(\geq 10 \text{ correct} \mid H_0) < .05 \)
Hypothesis Testing

• Why the “or more extreme” bit?
  • in larger data sets, even the most probable results are very unlikely (e.g., 500 coin flips)
  • with 50 coin flips, making 30 correct predictions has $p = .042$
  • if 30 is taken as evidence against $H_0$, then so should 31, or 32, all the way up to 50 correct
  • under $H_0$, how likely are we to get 30 or more correct? $p = .102$
Significance Test for $r$

• Section 9.10

• Population parameter $\rho$ null hypothesis: $\rho = 0$
  • even if $\rho$ is 0, it is likely that a (random) sample will produce $r \neq 0$

• Consider a small population of scores with $\rho = 0$
Significance Test for $r$

- Section 9.10

- Population parameter $\rho$ null hypothesis: $\rho = 0$
  - even if $\rho$ is 0, it is likely that a (random) sample will produce $r \neq 0$

- Consider a small population of scores with $\rho = 0$
Significance Test for $r$

- Population parameter $\rho$  null hypothesis: $\rho = 0$
  - is the observed value of $r$ among the least likely expected under the null hypothesis?
- what does the distribution of sample values of $r$ look like under the null hypothesis?
  - roughly normal, depending on sample size
  - as sample size increases, variability of $r$ values decreases

![Distribution Diagram](image-url)
Significance Test for $r$

- Table of critical values for $r$ (Table E.2 in textbook)
- depends on sample size and significance level
  - degrees of freedom: $df = N - 2$
- significance level usually .05 (least likely values)
  - two-tailed $p$ values

<table>
<thead>
<tr>
<th>$df$</th>
<th>$p = .10$</th>
<th>$p = .05$</th>
<th>$p = .025$</th>
<th>$p = .01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.805</td>
<td>0.878</td>
<td>0.924</td>
<td>0.959</td>
</tr>
<tr>
<td>4</td>
<td>0.729</td>
<td>0.811</td>
<td>0.868</td>
<td>0.917</td>
</tr>
<tr>
<td>5</td>
<td>0.669</td>
<td>0.755</td>
<td>0.817</td>
<td>0.875</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>60</td>
<td>0.211</td>
<td>0.250</td>
<td>0.285</td>
<td>0.325</td>
</tr>
<tr>
<td>120</td>
<td>0.150</td>
<td>0.178</td>
<td>0.203</td>
<td>0.232</td>
</tr>
<tr>
<td>200</td>
<td>0.116</td>
<td>0.138</td>
<td>0.158</td>
<td>0.181</td>
</tr>
<tr>
<td>500</td>
<td>0.073</td>
<td>0.088</td>
<td>0.100</td>
<td>0.115</td>
</tr>
</tbody>
</table>
Significance Test for $r$

- Reporting significance test for $r$
  - $N = 70 \quad r = .30$
  - $r(68) = .30, \quad p < .05$

<table>
<thead>
<tr>
<th>$df$</th>
<th>$p = .10$</th>
<th>$p = .05$</th>
<th>$p = .025$</th>
<th>$p = .01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.805</td>
<td>0.878</td>
<td>0.924</td>
<td>0.959</td>
</tr>
<tr>
<td>4</td>
<td>0.729</td>
<td>0.811</td>
<td>0.868</td>
<td>0.917</td>
</tr>
<tr>
<td>5</td>
<td>0.669</td>
<td>0.755</td>
<td>0.817</td>
<td>0.875</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>60</td>
<td>0.211</td>
<td>0.250</td>
<td>0.285</td>
<td>0.325</td>
</tr>
<tr>
<td>120</td>
<td>0.150</td>
<td>0.178</td>
<td>0.203</td>
<td>0.232</td>
</tr>
<tr>
<td>200</td>
<td>0.116</td>
<td>0.138</td>
<td>0.158</td>
<td>0.181</td>
</tr>
<tr>
<td>500</td>
<td>0.073</td>
<td>0.088</td>
<td>0.100</td>
<td>0.115</td>
</tr>
</tbody>
</table>
Significance Test for $r$

- **Using R** to test significance of $r$

```r
> dat=read.table(file.choose(new=T),header=T)
> plot(dat)
> head(dat)

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33 81</td>
</tr>
<tr>
<td>2</td>
<td>38 64</td>
</tr>
<tr>
<td>3</td>
<td>30 76</td>
</tr>
<tr>
<td>4</td>
<td>41 74</td>
</tr>
<tr>
<td>5</td>
<td>38 66</td>
</tr>
<tr>
<td>6</td>
<td>38 63</td>
</tr>
</tbody>
</table>
```
Significance Test for $r$

- Using R to test significance of $r$

```r
> cor.test(dat$X, dat$Y)

Pearson's product-moment correlation

data:  dat$X and dat$Y
t = -3.1535, df = 38, p-value = 0.003147
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:  
-0.6716504 -0.1677122
sample estimates:  
  cor
-0.4554282
```