P-value approach to H.T.
- based on computer output
  - no critical value is reported
  - instead, a "p-value" is reported

Critical Region

\[ H_1 : p > 20 \]

\[ \alpha = 0.025 \]

Right tail

Test statistic

\[ \text{stat value} = 1.31 \]

\[ p = 0.025 \]

"P-value" is area to the right
(because it is a right tail test)

\[ Z_d = 1.96, \quad p = 0.025 \]

If test stat = 2.14

\[ p = 0.0162 \]

If p-value is less than \( \alpha \), then reject \( H_0 \):

Otherwise, do not reject \( H_0 \):
If test stat = 1.81

Ch. 11-2  \{ counts in categories \}

Ch. 11-3  Each category has a "p"

Ch. 11-4  means for more than 2 treatments

Ch. 11-2  Multinomial  or  Goodness-of-fit

Is this die fair?

p555

<table>
<thead>
<tr>
<th>X</th>
<th>Frequency observed</th>
<th>Expected</th>
<th>((O-E)^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105</td>
<td>100</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>92</td>
<td>100</td>
<td>0.64</td>
</tr>
<tr>
<td>3</td>
<td>96</td>
<td>100</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>101</td>
<td>100</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>112</td>
<td>100</td>
<td>1.44</td>
</tr>
<tr>
<td>6</td>
<td>94</td>
<td>100</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>600</td>
<td></td>
<td>2.86</td>
</tr>
</tbody>
</table>

Test stat \(\sum \frac{(O-E)^2}{E}\)

\(H_0: \) die is fair

\(H_1: \) die is not fair

\(H_0: p_1 = \frac{1}{6}, p_2 = \frac{1}{6}, p_3 = \frac{1}{6}, p_4 = \frac{1}{6}, p_5 = 1, p_6 = 1\)

\(H_1: \) not \(H_0: \)
$\alpha = 0.05$ Right tail

Critical Region

Chi-square

d.f. = number of categories - 1 = $k-1$

$d.f. = 6 - 1$

\[ \chi^2 = 9.26 \]

Do not Reject $H_0$.

Quis 18 (2)

Example Exam Set p. 16 #4

Contingency Tables

\[ E = \frac{(\text{row total})(\text{cat. total})}{\text{grand total}} \]

<table>
<thead>
<tr>
<th></th>
<th>Toy</th>
<th># Kids</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3</td>
<td>14</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>16</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>30</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

\[ d.f. = (r-1)(c-1) = (2-1)(3-1) = 2 \]

<table>
<thead>
<tr>
<th>Gender</th>
<th>Toy</th>
<th># Kids</th>
<th></th>
<th>B</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Ball</td>
<td>40</td>
<td></td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>G</td>
<td>Ball</td>
<td>20</td>
<td></td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>Doll</td>
<td>5</td>
<td></td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>G</td>
<td>Doll</td>
<td>25</td>
<td></td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>Bell</td>
<td>5</td>
<td></td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>G</td>
<td>Bell</td>
<td>5</td>
<td></td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Test of Homogeneity

We test the claim that different populations have the same proportion of some characteristic.

Example:

$H_0$: People in the different age groups use the slang terms in the same proportions.

$H_1$: Independent, in $H_0$.

II-4

Analysis of Variance

p. 583

Treatment

Blood Pressure

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>\bar{X}</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td>112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>128</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>114</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\bar{X} = \frac{120 + 112 + 110 + 128 + 108 + 119 + 109 + 114}{8}$

$s^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$

$\sum (X - \bar{X})^2 = 216$

$n = 8$

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$

$H_1$: not $H_0$.

Assume: $\sigma^2 = \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = \sigma_6^2$

$\sigma^2$ homogeneous variances
Difference between treatments (groups)
Differences within treatments (groups)

Ex Exam Set p 17 # 6

\[
\begin{align*}
\bar{x}_1 - \bar{x}_2 & \quad \begin{array}{cc}
G_1 & G_2 \\
10 & 31 \\
9 & 30 \\
10 & 29 \\
10 & 29 \\
11 & 30 \\
\end{array} \\
\text{not random} \\
\end{align*}
\]

\[
\begin{array}{cc}
\text{Group} & \text{Group} \\
10 & 12 \\
30 & 31 \\
41 & 39 \\
22 & 25 \\
50 & 52 \\
\end{array}
\]

Variation between groups
Variation within groups

\[
S^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + (n_3-1)s_3^2 + \ldots + (n_k-1)s_k^2}{(n_1-1)+(n_2-1)+\ldots+(n_k-1)+(n_k-1)}
\]

\[
\sum_{i=1}^{k} n_i \left( x_i - \bar{x} \right)^2 \rightarrow \text{treatment sum of squares}
\]
Total Variation = Explained + Unexplained variation.

\[ \sum (x - \bar{x})^2 = \text{Total sum of squares} = \]

\[ \quad = \text{Treatment SS} + \text{Error SS} \]

between groups
within groups

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>Test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>(k - 1)</td>
<td>(SS(\text{treat}))</td>
<td>(\frac{SS(\text{treat})}{k-1})</td>
<td>(\frac{MS(\text{treat})}{MS(\text{error})})</td>
</tr>
<tr>
<td>Error</td>
<td>(N - K)</td>
<td>(SS(\text{error}))</td>
<td>(\frac{SS(\text{error})}{N-K})</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(N - 1)</td>
<td>(SS(\text{total}))</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\[ s_p^2 = \frac{SS(\text{error})}{N - K} \]

\[ SS(\text{treat}) = MS(\text{treat}) \times \text{d.f.} \]

p. 618  F Distribution (\(\alpha = 0.05\))

[Diagram of F distribution with critical value 3.2389]