This is two independent samples (dichotomous outcomes)

next time: correlation & regression

today: LN p. 1 - 243

Red: LN p. L - 214

if you have a red disc see, this week only see course web page about which section to go to (tomorrow is a holiday)

case study: sudden oak death

<table>
<thead>
<tr>
<th>where</th>
<th>sample</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>1</td>
<td>9/265 = 3.4%</td>
</tr>
<tr>
<td>OR</td>
<td>2</td>
<td>20/281 = 7.1%</td>
</tr>
</tbody>
</table>

7.1% - 3.4% = 3.4% = 3.4%

so d.n.e i. or is 109% bigger than s.o.d.

rok i.e. CA

EV of \( p_i \) = \( E_{\text{exp}}(p_i) = \hat{p}_i \)

\[
E_{\text{exp}}(p_i) - p_i
\]
\[
\text{SE (\(\hat{p}_1\))} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} = \frac{0.0111}{\sqrt{265}} = 0.1111 \approx 1.1\%
\]

\[
\text{SE (\(\hat{p}_2\))} = \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = \sqrt{\frac{(0.071)(0.929)}{281}} = 0.0153 \approx 1.5\%
\]

\[
-1.96 (\hat{p}_1 - \hat{p}_2) + 1.96
\]

\[
-1.96 (\hat{p}_1 - \hat{p}_2)
\]
\[ SE(\hat{p}_1 - \hat{p}_2) = \sqrt{SE(\hat{p}_1)^2 + SE(\hat{p}_2)^2} \]

\[ SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\left( \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} \right)^2 + \left( \frac{\hat{p}_2(1-\hat{p}_2)}{n_2} \right)^2} \]

\[ SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \]
\[
(\hat{p}_1 - \hat{p}_2) \pm 1.96 \cdot \text{SE}(\hat{p}_1 - \hat{p}_2)
\]

\[-3.7\% \pm (2)(1.9\%) = -3.7\% \pm 3.8\%
\]

\[
95\% \text{ CI for } (\hat{p}_1 - \hat{p}_2)
\]

\[
\left( \frac{-1}{1} \right)
\]

\[-7.5\% \quad -3.7\% \quad +0.1\%
\]

0\% is in 95\% CI, so not strictly speaking statistically, but evidence tending toward conclusion that this diff is real

scattered plot

elliptical

shape

bivariate

use x to predict y
facts about $r$

1. $r$ is a pure number without units

2. $1 + r = 1$

$r = -1$

$r = 1$