BE640 - Intermediate Biostatistics

Topic 3 – Discrete Distributions
Software: Stata v 7

Fisher’s Exact Test


Smith, Delgado and Rutledge (1976) report data on ovarian carcinoma. Individuals had different numbers of courses of chemotherapy. The 5-year survival data for those with 1-4 and 10 or more courses of chemotherapy are:

<table>
<thead>
<tr>
<th>Courses</th>
<th>Five Year Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dead</td>
</tr>
<tr>
<td>1-4</td>
<td>21</td>
</tr>
<tr>
<td>≥ 10</td>
<td>2</td>
</tr>
</tbody>
</table>

Using Fisher’s Exact test, is there a statistically significant association (p<.05) in this table?

There are multiple ways to do this. Three of them are

1. the `tabi` instruction to read in the data row by row and the `exact` option.
   This yields p-value only; there are no additional results and no fancy output.

2. the `tabulate` instruction with an existing data set and the `exact` option.
   This yields p-value only; there are no additional results but output can be prettier.

3. the `cs` instruction with an existing data set and options `FREQ=`, `OR` and `EXACT`.
   This yields p-values, estimated OR and the output can look nice.
1 - \texttt{tabi} instruction (data is read in row by row) and the \texttt{exact} option.

```
.tabi 21 2 \\ 8, exact
```

<table>
<thead>
<tr>
<th>row</th>
<th>col</th>
<th>1</th>
<th>2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
<td>2</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>8</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>10</td>
<td></td>
<td>33</td>
</tr>
</tbody>
</table>

Fisher's exact = 0.000
1-sided Fisher's exact = 0.000

2 - \texttt{tabulate} instruction (data already exists) and the \texttt{exact} option.

Enter your data into the \texttt{DATA EDITOR}

```
tally[1] = 21
```

```
courses | vital | tally
---------|-------|------
        1 | 1     | 21   
        2 | 1     | 2    
        3 | 0     | 2    
        4 | 0     | 8    
```
If you like, assign labels to the variable names by double clicking on the column headings. Then exit the Data Editor.

![Stata Variable Information]

Make your output readable

```
. * Dictionary of variable value labels
. label define coursesf 1 "1=1-4" 0 "0=10+
. label define vitalf 1 "1=dead" 0 "0=alive"

. * Use labels from dictionary
. label values courses coursesf
. label values vital vitalf
```

Two approaches to get Fisher’s Exact and Empirical Odds Ratio

- 1 - Tabulate
- 2 - CS

**Tabulate** with option **exact** will give you Fisher’s Exact Test but no measures of association

```
. tabulate courses vital [fweight=tally], exact

<table>
<thead>
<tr>
<th>Chemothera</th>
<th></th>
<th>vital status</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>O=alive</td>
<td>I=dead</td>
<td></td>
</tr>
<tr>
<td>0=10+</td>
<td>8</td>
<td>2</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>1=1-4</td>
<td>2</td>
<td>21</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>23</td>
<td></td>
<td>33</td>
</tr>
</tbody>
</table>

Fisher's exact = 0.000
1-sided Fisher's exact = 0.000
```
3 - cs instruction (data already exists) and options `FREQ=`, `OR` and `EXACT`

```
.cs courses vital [freq=tally], or exact

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Unexposed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>21</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>Noncases</td>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>10</td>
<td>33</td>
</tr>
</tbody>
</table>

Risk | .9130435 | .2 | .6969697

Risk difference | .7130435 | .4396867 | .9864002
Risk ratio | 4.565217 | 1.313213 | 15.8704
Attr. frac. ex. | .7809524 | .2385087 | .9369896
Attr. frac. pop | .7130435 |
Odds ratio | 42 | 5.530036 | 321.0117 (Cornfield)

1-sided Fisher's exact P = 0.0001
2-sided Fisher's exact P = 0.0001