

**BE640 - Intermediate Biostatistics
Computer Illustration**

**Topic 3 – Discrete Distributions
Software: Stata v 7**

Fisher's Exact Test

Source: Fisher LD and Van Belle G. Biostatistics: A Methodology for the Health Sciences. New York: Wiley, 1993. Chapter 6 problem 5, page 232.

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:

Courses	<u>Five Year Status</u>	
	<u>Dead</u>	<u>Alive</u>
1-4	21	2
≥ 10	2	8

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- **tabi** instruction (data is read in row by row) and the **exact** option.

.tabi 21 2 \2 8, exact

row	col		Total
	1	2	
1	21	2	23
2	2	8	10
Total	23	10	33

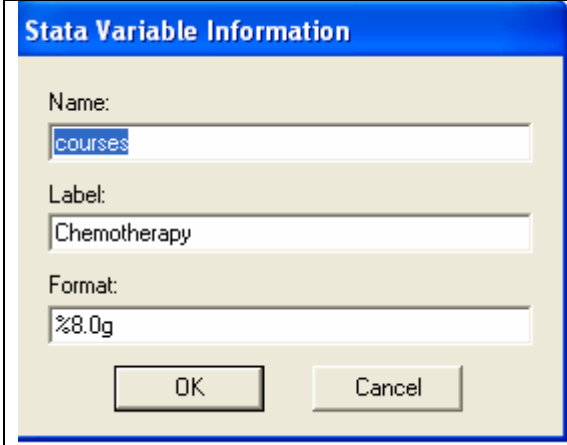
Fisher's exact = 0.000
 1-sided Fisher's exact = 0.000

2- **tabulate** instruction (data already exists) and the **exact** option.

Enter your data into the **DATA EDITOR**

Stata Editor				
<input type="button" value="Preserve"/> <input type="button" value="Restore"/> <input type="button" value="Sort"/> <input type="button" value="<<"/> <input type="button" value=">>"/> <input type="button" value="Hide"/> <input type="button" value="Delete..."/>				
tally[1] = 21				
	courses	vital	tally	
1	1	1	21	
2	1	0	2	
3	0	1	2	
4	0	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

Name:

Label:

Format:

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 coursesf coursesf
. label values vitalf 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
```

Chemothera py	vital status		Total
	0=alive	1=dead	
0=10+	8	2	10
1=1-4	2	21	23
Total	10	23	33

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
```

	vital status			
	Exposed	Unexposed	Total	
Cases	21	2	23	
Noncases	2	8	10	
Total	23	10	33	
Risk	.9130435	.2	.6969697	
	Point estimate		[95% Conf. Interval]	
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			