BE640 - Intermediate Biostatistics
Computer Illustration

Topic 3 – Discrete Distributions
Software: SAS

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:

	Five Year Status	
Courses	**Dead**	**Alive**
1-4	21	2
≥ 10	2	8

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

Suggestions -

(1) Use 0/1 convention for coding alive/dead with 1 = dead
(2) Use 0/1 convention for coding exposure with 1 = exposed
(3) Use the WEIGHT instruction to enter the cell counts.

```
options nocenter nodate;
proc format;                                    Proc FORMAT creates dictionary of labels.
  value coursef 1="1=1-4 courses"
                0="0=10 plus";
  value vitalf 1="1=dead"
               0="0=alive";
data temp;
  input courses vital count;
  cards;
  1 1 21
  1 0 2
  0 1 2
  0 0 8
  ;
 run;
 proc freq data=temp;
   tables courses*vital/exact cmh;
   weight count;
   format courses coursef. vital vitalf.;    FORMAT utilizes labels for readable output.
 run;
```

You should see the following.

```
The SAS System

The FREQ Procedure
Table of courses by vital

courses        vital

Frequency
Percent
Row Pct
Col Pct        |0=alive |1=dead  |  Total
---------------+--------+--------+--------
0=10 plus      |      8 |      2 |     10
               |  24.24 |   6.06 |  30.30
               |  80.00 |  20.00 |
               |  80.00 |   8.70 |
---------------+--------+--------+--------
1=1-4 courses  |      2 |     21 |     23
               |   6.06 |  63.64 |  69.70
               |   8.70 |  91.30 |
               |  20.00 |  91.30 |
---------------+--------+--------+--------
Total               10       23       33
                 30.30    69.70   100.00


Statistics for Table of courses by vital

Statistic                     DF       Value      Prob
------------------------------------------------------
Chi-Square                     1     16.7782    <.0001
Likelihood Ratio Chi-Square    1     16.8868    <.0001
Continuity Adj. Chi-Square     1     13.5720    0.0002
Mantel-Haenszel Chi-Square     1     16.2698    <.0001
Phi Coefficient                      0.7130
Contingency Coefficient              0.5806
Cramer's V                           0.7130

 WARNING: 25% of the cells have expected counts less
          than 5. Chi-Square may not be a valid test.


       Fisher's Exact Test
----------------------------------
Cell (1,1) Frequency (F)          8
Left-sided Pr <= F           1.0000
Right-sided Pr >= F      1.255E-04

Table Probability (P)    1.230E-04
Two-sided Pr <= P        1.255E-04

Sample Size = 33
```

SAS gives you the 2x2 table so that you can check on its formulation of risk.

SAS will do an overall chi square test. Beware, the chi square is an approximate test that is not always valid.

WARNING: SAS is telling us that the overall chi square test may not be valid because of low cell frequencies.

EXACT option will give you the Fisher Exact test. It's limitation is that it is computer intensive when the cell frequencies get big. Here, however, it's just fine.

Two sided p-value is < .0001

The FREQ Procedure

Summary Statistics for courses by vital

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	16.2698	<.0001
2	Row Mean Scores Differ	1	16.2698	<.0001
3	General Association	1	16.2698	<.0001

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	Mantel-Haenszel	42.0000	5.0293	350.7474
	Logit	42.0000	5.0293	350.7474
Cohort (Col1 Risk)	Mantel-Haenszel	9.2000	2.3612	35.8469
	Logit	9.2000	2.3612	35.8469
Cohort (Col2 Risk)	Mantel-Haenszel	0.2190	0.0630	0.7615
	Logit	0.2190	0.0630	0.7615

Total Sample Size = 33

CMH option gives you estimates of relative risk for 2 study designs plus the estimated odds ratio. Here, the estimated OR = 42.