Chapter 2: The 2 × 2 Table

1. Introduction

- Simple random samples from two groups that yield two independent binomial distributions for a binary response
- Simple random sample from one group that yields a single multinomial distribution for the cross-classification of two binary responses
- Randomized assignment of patients to two equivalent treatments, resulting in the hyper-geometric distribution.
- Counts of events from independent Poisson processes

2. Chi-square Statistics

Treatment	Fav	Unfav	Total	
Placebo	$n_{11}=16$	$n_{12}=48$	$n_{1+}=64$	$p_1=(n_{11}/n_{1+})=0.25$
Test	$n_{21}=40$	$n_{22}=20$	$n_{2+}=60$	$p_2=(n_{21}/n_{2+})=0.67$
Total	$n_{+1}=56$	$n_{+2}=68$	$n=124$	$\overline{p}=(n_{+1}/n)=0.45$

H_0: There is no association between treatment and response (treatments have equal effects)

Under H_0, $\Pr\{n_{ij}\}=\dfrac{n_{1+}!n_{2+}!n_{+1}!n_{+2}!}{n!n_{11}!n_{12}!n_{21}!n_{22}!}$

$$E\{n_{ij}\,|H_0\}=\frac{n_{i+}n_{+j}}{n}=m_{ij} \qquad V\left\{n_{ij}\,|H_0\right\}=\frac{n_{1+}n_{2+}n_{+1}n_{+2}}{n^2(n-1)}=v_{ij}$$

Randomization chi-square $Q = \dfrac{(n_{11} - m_{11})^2}{v_{11}} = 21.53$

Pearson chi-square $Q_P = \sum_{i=1}^{2}\sum_{j=1}^{2} \dfrac{(n_{ij} - m_{ij})^2}{m_{ij}} = \dfrac{n}{(n-1)} Q = 21.71$

Both Q and Q_P have p-values < 0.001

Note: $(n_{11} - m_{11}) = \dfrac{(n_{11}n_{22} - n_{12}n_{21})}{n} = \dfrac{n_{1+}n_{2+}}{n}(p_1 - p_2)$

$$Q_P = (p_1 - p_2)^2 \Big/ \left\{ \left(\frac{1}{n_{1+}} + \frac{1}{n_{2+}} \right) \bar{p}(1-\bar{p}) \right\} \qquad \bar{p} = n_{+1}/n$$

	F	*U*	Total
P	16	48	64
T	40	20	60
Total	56	68	124

P : 25% favorable

T : 67% favorable

```
data respire;
   input  treat  $  outcome  $  count;
   datalines;
   placebo   f  16
   placebo   u  48
   test      f  40
   test      u  20
;
```

PROC FREQ

- Nonparametric procedure used to describe distributions and associations of categorical data through contingency tables

- Testing
  Produces Chi-square test to analyze 2×2 table, Mantel-Haenszel test and Fisher's test for 2×2 table,

- Measures of Association
  Produces risk difference, odds ratio, and relative risk with approximate confidence intervals for a single 2×2 table

  - McNemar's test

```
PROC FREQ options;
   OUTPUT <OUT= SAS-data-set><output-statistic-list>;
   TABLES requests / options;
   WEIGHT variable;
   EXACT statistic-keywords;
   BY variable-list;
```

- Expanded capabilities for exact *p*-values and confidence intervals via the EXACT statement
    - Exact confidence interval for odds ratio from a 2×2 table
    - Exact confidence interval for risk difference (SAS 9.2 and 9.3) from a 2×2 table
    - Exact results for McNemar's test and the binomial proportion test

```
proc freq;
   weight count;
   tables treat*outcome / chisq;
run;
```

Table of treat by outcome

treat	outcome		
Frequency Percent Row Pct Col Pct	f	u	Total
placebo	16 12.90 25.00 28.57	48 38.71 75.00 70.59	64 51.61
test	40 32.26 66.67 71.43	20 16.13 33.33 29.41	60 48.39
Total	56 45.16	68 54.84	124 100.00

Statistics for Table of treat by outcome

Statistic	DF	Value	Prob
Chi-Square	1	21.7087	<.0001
Likelihood Ratio Chi-Square	1	22.3768	<.0001
Continuity Adj. Chi-Square	1	20.0589	<.0001
Mantel-Haenszel Chi-Square	1	21.5336	<.0001
Phi Coefficient		-0.4184	
Contingency Coefficient		0.3860	
Cramer's V		-0.4184	

Fisher's Exact Test

Cell (1,1) Frequency (F)	16
Left-sided Pr <= F	2.838E-06
Right-sided Pr >= F	1.0000
Table Probability (P)	2.397E-06
Two-sided Pr <= P	4.754E-06

Sample Size = 124

$Q_P = 21.71, p < 0.001$

$Q = 21.53, p < 0.001$

3. Exact Tests

Treatment	Fav	Unfav	Total
Test	10	2	12
Control	2	4	6
Total	12	6	18

$$\Pr\{n_{ij}\} = \frac{n_{1+}!n_{2+}!n_{+1}!n_{+2}!}{n!n_{11}!n_{12}!n_{21}!n_{22}!}$$

Table Cell

(1,1)	(1,2)	(2,1)	(2,2)	Probabilities
12	0	0	6	0.0001
11	1	1	5	0.0039
10	2	2	4	0.0533
9	3	3	3	0.2370
8	4	4	2	0.4000
7	5	5	1	0.2560
6	6	6	0	0.0498

One-sided *p*-value:

p = 0.0533 + 0.0039 + 0.0001= 0.0573

Two-sided *p*-value:

p = 0.0533 + 0.0039 + 0.0001 + 0.0498 = 0.1071

```
data severe;
   input  treat  $  outcome  $  count;
   datalines;
   test             f      10
   test             u       2
   control          f       2
   control          u       4
;

proc freq order=data;
   weight count;
   tables treat*outcome / chisq nocol;
run;
```

TREAT	OUTCOME		
Frequency Percent Row Pct	f	u	Total
test	10 55.56 83.33	2 11.11 16.67	12 66.67
control	2 11.11 33.33	4 22.22 66.67	6 33.33
Total	12 66.67	6 33.33	18 100.00

Statistics for Table of treat by outcome

Statistic	DF	Value	Prob
Chi-Square	1	4.5000	0.0339
Likelihood Ratio Chi-Square	1	4.4629	0.0346
Continuity Adj. Chi-Square	1	2.5313	0.1116
Mantel-Haenszel Chi-Square	1	4.2500	0.0393
Phi Coefficient		0.5000	
Contingency Coefficient		0.4472	
Cramer's V		0.5000	

WARNING: 75% 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)	10
Left-sided Pr <= F	0.9961
Right-sided Pr >= F	0.0573
Table Probability (P)	0.0533
Two-sided Pr <= P	0.1070

Sample Size = 18

Exact *p*-values for Chi-Square Statistics

```
proc freq order=data;
   weight count;
   tables treat*outcome / chisq nocol;
   exact chisq;
run;
```

Pearson Chi-Square Test	
Chi-Square	4.5000
DF	1
Asymptotic Pr > ChiSq	0.0339
Exact Pr >= ChiSq	0.1070

Likelihood Ratio Chi-Square Test	
Chi-Square	4.4629
DF	1
Asymptotic Pr > ChiSq	0.0346
Exact Pr >= ChiSq	0.1070

Mantel-Haenszel Chi-Square Test	
Chi-Square	4.2500
DF	1
Asymptotic Pr > ChiSq	0.0393
Exact Pr >= ChiSq	0.1070

Consider a very large, randomized field study to compare failure rates for a new vaccine versus a control condition.

Suppose the data are

	Fail	Success	Total
Vaccine	n_1	$N_1 - n_1$	N_1
Control	n_2	$N_2 - n_2$	N_2
Total	n	$N - n$	N

Note that N_1, N_2 are very large; but n_1, n_2 are very small; e.g., $n \leq 20$ and $N \geq 1000$. Consider Fisher's test probability function for n_1

$$\Pr\{n\} = \frac{N_1!N_2!(N-n)!n!}{n_1!n_2!(N_1-n_1)!(N_2-n_2)!N!}$$

$$\approx \frac{n!}{n_1!n_2!}\left(\frac{N_1}{N}\right)^{n_1}\left(\frac{N_2}{N}\right)^{n_2}$$

Since $N_1 \gg n_1$, $N_2 \gg n_2$, $\Pr\{n_1\}$ is binomial $\text{Bin}\left(n, \frac{N_1}{N}\right)$.

Thus, the left-hand (lower) tail *p*-value for Fisher's test in this case is

$$p = \sum_{j=0}^{n_1} \frac{n!}{j!(n-j)!} \left(\frac{N_1}{N}\right)^j \left(\frac{N_2}{N}\right)^{n-j}$$

For moderately large *n* so that $\frac{nN_1}{N}, \frac{nN_2}{N} \geq 10$, then n_1 is

Approximately normal $N\left(\frac{nN_1}{N}, \frac{nN_1N_2}{N^2}\right)$.

Example

	Fail	Success	Total
Vaccine	1	19,999	20,000
Control	5	9,995	10,000
Total	6	29,994	30,000

$$p = \binom{6}{0}\left(\frac{1}{3}\right)^6 + \binom{6}{1}\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5 = \frac{13}{3^6} = \frac{13}{729} \approx 0.02$$

Confidence Intervals For a Binomial Proportion

Given a random variable Y that is distributed Bin(n,P), the following confidence intervals apply:

- If $\min\{y,(n-y)\} \geq 15,$ use $p \pm z_{\alpha/2}\sqrt{p(1-p)/n},$ where $p = y/n$

- If $8 \leq \min\{y,(n-y)\} \leq 15,$ use

$$p \pm \left\{ z_{\alpha/2}\sqrt{p(1-p)/n} + \frac{1}{2n} \right\}$$

- If $5 \le \min\{y, (n-y)\} \le 8$, use solutions of quadratic equations (at 0.95, like $p^* = (y+2)/(n+4)$ for top)

- If $0 \le \min\{y, (n-y)\} \le 5$, use solutions to:

$$\sum_{x \ge y} \binom{n}{x} P_L^x (1-P_L)^{n-x} = \frac{\alpha}{2} \uparrow P$$

$$\sum_{x \le y} \binom{n}{x} P_U^x (1-P_U)^{n-x} = \frac{\alpha}{2} \downarrow P,$$

For the quadratic equation method, suppose Y is Bin(n, P) and that n is large; e.g., $n \geq 20$ and

$nP,\ n(1-P) \geq 5$

Then $(Y-nP)/\{nP(1-P)\}^{\frac{1}{2}}$ is approximately N(0,1). With a continuity correction, any value of P such that

$$\{|(Y-nP)|-0.5\}^2 / nP(1-P) \geq Z^2_{\alpha/2}$$

where $Z_{a/2}$ is the $100(1-a/2)$ percentile of N(0, 1), would be rejected by two-sided $p \leq \alpha$ in a corresponding hypothesis test. The $100(1-\alpha)\%$ confidence interval contains all other values of P.

The limits are identified by solving:

$$(Y-nP)^2 - |Y-nP| + 0.25 = Z^2_{\alpha/2} nP(1-P).$$

The upper and lower limits P_U and P_L of this interval are

$$P_U = \frac{\left(\frac{y}{n} + \frac{1}{2n} + \frac{Z^2_{\alpha/2}}{2n}\right) + Z_{\alpha/2}\sqrt{\frac{Z^2_{\alpha/2}}{4n^2} + \frac{1}{4n^2}\left(2 - \frac{1}{n}\right) + \frac{1}{4n^2}4Y\left(1 - \frac{y}{n} - \frac{1}{n}\right)}}{\left(1 + \frac{Z^2_{\alpha/2}}{n}\right)}$$

$$\text{As } n \uparrow \infty, P_U \to \frac{Y}{n} + Z_{\alpha/2}\sqrt{\frac{Y}{n^2}\left(1 - \frac{Y}{n}\right)} = p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}, \text{ where } p = \left(\frac{Y}{n}\right)$$

$$P_L = \frac{\left(\frac{Y}{n} - \frac{1}{2n} + \frac{Z^2_{\alpha/2}}{2n}\right) - Z_{\alpha/2}\sqrt{\frac{Z^2_{\alpha/2}}{4n^2} - \frac{1}{4n^2}\left(2 + \frac{1}{n}\right) + \frac{1}{4n^2}4Y\left(1 - \frac{Y}{n} + \frac{1}{n}\right)}}{\left(1 + \frac{Z^2_{\alpha/2}}{n}\right)}$$

$$\text{As } n \uparrow \infty, P_L \to \frac{Y}{n} - Z_{\alpha/2}\sqrt{\frac{Y}{n^2}\left(1 - \frac{Y}{n}\right)} = p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}, \text{ where } p = \left(\frac{Y}{n}\right)$$

4. Difference in Proportions

Group	Yes	No	Total	Proportion Yes
1	n_{11}	n_{12}	n_{1+}	$p_1 = n_{11}/n_{1+}$
2	n_{21}	n_{22}	n_{2+}	$p_2 = n_{21}/n_{2+}$
Total	n_{+1}	n_{+2}	n	$d = p_1 - p_2$

Source (SDK, 4)

Assume Groups 1 and 2 have independent simple random samples so that independent binomial distributions apply

$$E\{p_1 - p_2\} = \pi_1 - \pi_2$$

$$V\{p_1 - p_2\} = \frac{\pi_1(1-\pi_1)}{n_{1+}} + \frac{\pi_2(1-\pi_2)}{n_{2+}}$$

with consistent estimate: $v_d = \frac{p_1(1-p_1)}{n_{1+}} + \frac{p_2(1-p_2)}{n_{2+}}$

A 100(1 – α) % confidence interval for $\pi_1 - \pi_2$ is written:

$$d \pm \left\{ z_{\alpha/2} \sqrt{v_d} + \frac{1}{2} \left[\frac{1}{n_{1+}} + \frac{1}{n_{2+}} \right] \right\}$$

$$\tilde{v}_d = \frac{p_1(1-p_1)}{n_{1+}-1} + \frac{p_2(1-p_2)}{n_{2+}-1}$$

is an unbiased estimator of the variance of *d* and can replace v_d in the confidence interval.

Pearson correlation coefficient:

$$r = \left\{ \left(n_{11} - \frac{n_{1+} n_{+1}}{n} \right) \Big/ \left[\left(n_{1+} - \frac{n_{1+}^2}{n} \right) \left(n_{+1} - \frac{n_{+1}^2}{n} \right) \right]^{1/2} \right\}$$

$$= \sqrt{Q_P / n}$$

We can use the quadratic equation method for a confidence interval on the difference of proportions. We plug in the quadratic solutions P_{U1}, P_{L1} for group 1 and P_{U2}, P_{L2} for group 2 into the formulas

$$L = (p_1 - p_2) - \sqrt{(p_1 - P_{L1})^2 + (P_{U2} - p_2)^2}$$

$$U = (p_1 - p_2) + \sqrt{(P_{U1} - p_1)^2 + (p_2 - P_{L2})^2}$$

See slide 20 for continuity corrected $P_{L1}, P_{U1}, P_{L2}, P_{U2}$. See page 30 of text for versions without continuity correction.

- RiskDiffCol1 table produces the difference for column 1 of the frequency table.
- Note that the RISKDIFF option does not use a continuity correction and has n's rather than $(n-1)$'s in the denominator of v_d and so may have lower than nominal (e.g., 95%) confidence level. The RISKDIFF (CORRECT) option does incorporate a continuity correction, but not $(n-1)$'s in the denominator or v_d.

```
data respire2;
   input treat $ outcome $ count @@;
   datalines;
test    f 40 test    u 20
placebo f 16 placebo u 48
;
ods select RiskDiffCol1 Measures;
proc freq order=data;
   weight count;
   tables treat*outcome / riskdiff (correct) measures;
run;
```

Results for Difference in Proportions

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.6667	0.0609	0.5391	0.7943	0.5331	0.7831
Row 2	0.2500	0.0541	0.1361	0.3639	0.1502	0.3740
Total	0.4516	0.0447	0.3600	0.5432	0.3621	0.5435
Difference	0.4167	0.0814	0.2409	0.5924		

Difference is (Row 1 - Row 2)
The asymptotic confidence limits include a continuity correction.

- Alternatives: Miettinen-Nurminem interval or corrected Wald interval
  - M-N might be preferred when cell count size is marginal

```
proc freq order=data;
   weight count;
   tables treat*outcome/riskdiff(cl=(wald mn) correct) measures;
run;
```

Confidence Limits for the Proportion (Risk) Difference
Column 1 (outcome = f)
Proportion Difference = 0.4167

Type	95% Confidence Limits	
Miettinen-Nurminen	0.2460	0.5627
Wald (Corrected)	0.2409	0.5924

- The Miettinen-Nurminen interval is obtained by identifying values of $\Delta = p_1 - p_2$ which produce values for Q_P which are ≤ 3.84 (for a 95% interval)

•Output shows a 95% M-N interval for Δ as (0.2460, 0.5627). The Wald continuity corrected interval is (0.2409, 0.5924). We can calculate values of Q_P for values of Δ around the M-N limits and show they are close to 3.84.

Lower Limit		**Upper Limit**	
Δ_L	Q_P	Δ_U	Q_P
0.2410	4.0618	0.6000	6.4616
0.2425	3.9957	0.5925	5.8626
0.2435	3.9520	0.5850	5.3018
0.2450	3.8868	0.5750	4.6102
0.2455	3.8651	0.5675	4.1315
0.2460	3.8436	0.5650	3.9792
0.2465	3.8221	0.5625	3.8306

0.2460 →

← 0.5627

- Alternatives: Newcombe hybrid score interval or exact interval
  - Newcombe method is asymptotic, but works well for small sample sizes (cell counts smaller than 8)
- Example: Severe infection treatment outcomes

Treatment	Fav	Unfav	Total
Test	10	2	12
Control	2	4	6
Total	12	6	18

- Newcombe method incorporating Wilson score confidence limits for Groups 1 and 2

$$L = (p_1 - p_2) - \sqrt{(p_1 - P_{L1})^2 + (P_{U2} - p_2)^2}$$

$$U = (p_1 - p_2) + \sqrt{(P_{U1} - p_1)^2 + (p_2 - P_{L2})^2}$$

See slide 20 for continuity corrected P_{L1}, P_{U1}, P_{L2}, P_{U2}. See page 30 of text for versions without continuity correction.

Exact Confidence Interval for Difference in Proportions

- SAS versions 9.2 and 9.3 contain an option for computing an exact confidence interval for the risk difference. This option is not available in version 9.1.3.

- Method is conservative in small samples. Coverage of the confidence interval is at least $100\times(1-\alpha)$ % (Agresti, 2003)

- Include EXACT RISKDIFF statement in PROC FREQ to calculate exact interval or request option (CL=EXACT). CORRECT option still needed to request continuity correction.

```
proc freq order=data data=severe;
   weight count;
   tables treat*outcome/riskdiff(cl=(wald newcombe exact) correct);
run;
```

Confidence Limits for the Proportion (Risk) Difference
Column 1 (outcome = f)
Proportion Difference = 0.5000

Type	95% Confidence Limits	
Exact	-0.0296	0.8813
Newcombe Score (Corrected)	-0.0352	0.8059
Wald (Corrected)	-0.0571	1.0000

Agresti's method:

Add 2 to counts for a single proportion

Add 1 to counts for difference of proportions

```
data agresti;
  set respire;
  count2 = count+2;
  count1 = count+1;

proc freq order=data data=agresti;
  weight count2;
  tables treat*outcome /  riskdiff ;

proc freq order=data data=agresti;
  weight count1;
  tables treat*outcome /  riskdiff ;
run;
```

Adding 2 to cell counts for single proportions:

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	**0.7500**	**0.1083**	**0.5378**	**0.9622**	**0.4762**	**0.9273**
Row 2	**0.4000**	**0.1549**	**0.0964**	**0.7036**	**0.1216**	**0.7376**
Total	0.6154	0.0954	0.4284	0.8024	0.4057	0.7977
Difference	0.3500	0.1890	-0.0204	0.7204		

Adding 1 to cell counts for difference of proportions:

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.7857	0.1097	0.5708	1.0000	0.4920	0.9534
Row 2	0.3750	0.1712	0.0395	0.7105	0.0852	0.7551
Total	0.6364	0.1026	0.4354	0.8374	0.4066	0.8280
Difference	**0.4107**	**0.2033**	**0.0123**	**0.8091**		

Difference is (Row 1 - Row 2)

5. Odds Ratio and Relative Risk

• The odds ratio compares the odds of the yes proportion for Group 1 to the odds of the yes proportion for Group 2.

$$OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

$$f = \log\{OR\} = \log\left\{\frac{p_1(1-p_2)}{p_2(1-p_1)}\right\}$$

$$= \log\{p_1/(1-p_1)\} - \log\{p_2/(1-p_2)\}$$

$$v_f \approx \left\{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}\right\}$$

A 100(1 – α) % confidence interval for OR can be written

$$\exp\left(f \pm z_{\alpha/2}\sqrt{v_f}\right)$$

- The relative risk is the risk of developing a particular condition for one group compared to another group.

$$RR = \frac{p_1}{p_2} = OR \times \frac{\{1+(n_{21}/n_{22})\}}{\{1+(n_{11}/n_{12})\}}$$

In order to produce chi-square statistics, odds ratios, and relative risk measures for the following data, the ALL option is used. This has the same action as specifying both the CHISQ and the MEASURES options (and the CMH option, discussed in Ch. 3)

Example 1

Consider the data from a study about how general daily stress affects one's opinions on a proposed new health policy. The data are cross-sectional.

Stress	Favorable	Unfavorable	Total
Low	48	12	60
High	96	94	190

SAS Code to enter data:

```
data stress;
input stress $ outcome $ count;
datalines;
low f 48
low u 12
high f 96
high u 94
;
```

The MEASURES option in PROC FREQ tells SAS to output information about the odds ratio and its confidence limits:

```
proc freq order=data;
       weight count;
       tables stress*outcome / chisq measures nocol nopct;
run;
```

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	3.9167	1.9575	7.8366
Cohort (Col1 Risk)	1.5833	1.3104	1.9131
Cohort (Col2 Risk)	0.4043	0.2389	0.6841

The estimate of the odds ratio is 3.92, with associated 95% CI of (1.96, 7.84). The odds of a favorable response are roughly 4 times as high as for those persons with low stress.

Example 2

Treatment	Yes	No	Total
Test	29	16	45
Placebo	14	31	45

```
data respire;
   input treat $ outcome $ count;
   datalines;
test          yes  29
test          no   16
placebo       yes  14
placebo       no   31
;

proc freq order=data;
    weight count;
   tables treat*outcome / all nocol nopct;
run;
```

Output 2.15 Table Statistics

Statistics for Table of treat by outcome

Statistic	DF	Value	Prob
Chi-Square	1	10.0198	0.0015
Likelihood Ratio Chi-Square	1	10.2162	0.0014
Continuity Adj. Chi-Square	1	8.7284	0.0031
Mantel-Haenszel Chi-Square	1	9.9085	0.0016
Phi Coefficient		0.3337	
Contingency Coefficient		0.3165	
Cramer's V		0.3337	

Fisher's Exact Test

Cell (1,1) Frequency (F)	29
Left-sided Pr <= F	0.9997
Right-sided Pr >= F	0.0015
Table Probability (P)	0.0011
Two-sided Pr <= P	0.0029

Output 2.17 Odds Ratio and Relative Risk

Estimates of the Relative Risk (Row1/Row2)

Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	4.013	1.668	9.656
Cohort (Col1 Risk)	2.071	1.274	3.368
Cohort (Col2 Risk)	0.516	0.333	0.801

Exact Confidence Limits for the Odds Ratio

• Odds ratios are used as a measure of association but the usual asymptotic confidence limits would not be appropriate when the table is sparse.

• One can obtain exact confidence limits for the odds ratio by using the non-central hypergeometric distribution

Example:

Treatment	Favorable	Unfavorable
Test	10	2
Control	2	4

Source (SDK, 23)

Asymptotic 95% Confidence Limits (1.03, 97.50)

Exact 95% Confidence Limits (0.69, 166.36)

• Exact results are more accurate.

```
data severe;
     input treat $ outcome $ count;
     datalines;
  Test    f 10
  Test    u 2
  Control f 2
  Control u 4
  ;
  proc freq order=data;
     weight count;
     tables treat*outcome / nocol;
     exact or;
  run;
```

Output 2.18 Odds Ratio (Case-Control Study)

Odds Ratio (Case-Control Study)	
Odds Ratio	10.0000
Asymptotic Conf Limits	
95% Lower Conf Limit	1.0256
95% Upper Conf Limit	97.5005
Exact Conf Limits	
95% Lower Conf Limit	0.6896
95% Upper Conf Limit	166.3562

6. Sensitivity and Specificity

- Sensitivity: true proportion of positive results that a test elicits when performed on subjects known to have the disease
  - Sensitivity = n_{11}/n_{1+} = Pr(Test + | disease present)

- Specificity: true proportion of negative results that a test elicits when performed on subjects known to be disease free
  - Specificity = n_{22}/n_{2+} = Pr(Test − | disease absent)

Example: Skin disease screening test results

Status	Test +	Test −	Total
Disease Present	52	8	60
Disease Absent	20	100	120

```
data screening;
   input disease $ outcome $ count @@;
   datalines;
present  +  52  present  -    8
absent   +  20  absent   -  100
run;

proc freq data=screening order=data;
   weight count;
   tables disease*outcome / riskdiff;
run;
```

- Sensitivity = Column 1 Estimate for Row 1

Statistics for Table of disease by outcome

Column 1 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.8667	0.0439	0.7807	0.9527	0.7541	0.9406
Row 2	0.1667	0.0340	0.1000	0.2333	0.1049	0.2456
Total	0.4000	0.0365	0.3284	0.4716	0.3278	0.4755
Difference	0.7000	0.0555	0.5912	0.8088		

Difference is (Row 1 - Row 2)

- Specificity = Column 2 Estimate for Row 2

Column 2 Risk Estimates

	Risk	ASE	(Asymptotic) 95% Confidence Limits		(Exact) 95% Confidence Limits	
Row 1	0.1333	0.0439	0.0473	0.2193	0.0594	0.2459
Row 2	0.8333	0.0340	0.7667	0.9000	0.7544	0.8951
Total	0.6000	0.0365	0.5284	0.6716	0.5245	0.6722
Difference	-0.7000	0.0555	-0.8088	-0.5912		

Difference is (Row 1 - Row 2)

- If you know the underlying percentage of subjects with and without the disease, you can use Bayes' Theorem to estimate the proportion of subjects with the disease among those who have a positive test.

- $\Pr(D \mid T) = \dfrac{\Pr(T, D)}{\Pr(T)}$

- If 15% of the population have the disease and 85% do not, we have

Status	Test +	Test −	Total
Disease Present	0.867(.15) = 0.130	0.133(.15) = 0.020	0.15
Disease Absent	0.167(.85) = 0.142	0.833(.85) = 0.708	0.85
Total	0.130+0.142=0.272	0.020+0.708=0.728	

- Pr(disease | Test +) = 0.130/0.272 = 0.478

- Pr(no disease | Test −) = 0.708/0.728 = 0.972

7. McNemar's Test

Response 1

Response 2	Yes	No	Total
Yes	n_{11}	n_{12}	n_{1+}
No	n_{21}	n_{22}	n_{2+}
Total	n_{+1}	n_{+2}	n

Are $p_1 = \frac{n_{+1}}{n}$ and $p_2 = \frac{n_{1+}}{n}$ the same?

$$Q_M = \frac{(n_{12} - n_{21})^2}{(n_{12} + n_{21})} \approx \chi^2(1)$$

$$Q_{M,C} = \frac{(|n_{12} - n_{21}| - 1)^2}{(n_{12} + n_{21})} \approx \chi^2(1)$$

Example

	Wife Approval		
Husband Approval	Yes	No	Total
Yes	20	5	25
No	10	10	20
Total	30	15	45

$$Q_M = \frac{(5-10)^2}{(5+10)} = 1.67$$

- McNemar's test may be computed with the FREQ procedure, including the AGREE option in the TABLE statement. The ODS SELECT statement is used to restrict the output to that test.

```
data approval;
   input hus_resp $ wif_resp $ count;
   datalines;
yes yes 20
yes no   5
no  yes  10
no  no   10
;

ods select McNemarsTest;
proc freq order=data;
   weight count;
   tables hus_resp*wif_resp/ agree;
   exact mcnem;
run;
```

Output 2.18 McNemar's Test

Statistics for Table of hus_resp by wif_resp

McNemar's Test

Statistic (S)	1.6667
DF	1
Asymptotic Pr > S	0.1967
Exact Pr >= S	0.3018

- Note that exact *p*-values are available for McNemar's test through the statement

```
exact mcnem;
```

8. Incidence Densities

- The 2×2 table can also represent incidence densities, in which you have counts of subjects who responded with an event vs. extent of exposure for that event

- Counts often follow the Poisson distribution

- Examples:
  - colony counts for bacteria or viruses
  - accidents or equipment failure
  - incidences for disease

- Often want to compute ratio of incidence densities

- Often want to compute ratio of incidence densities
- Example: Vaccine Study

Treatment	Events	Person Years
Vaccine	n_v	N_v
Control	n_c	N_c

Assume n_C and n_V have independent Poisson distributions with expected values $N_C\lambda_C$ and $N_V\lambda_V$, respectively.

The null hypothesis is $\lambda_1 = \lambda_2$ and n_V given $(n_V + n_C) = n$ is Bin$(n, N_V/N)$.

Let $n_C =$ number with disease for control

and $n_V =$ number with disease for vaccine

Assume:
- n_C is Poisson$(N_C\lambda_C)$
- n_V is Poisson$(N_V\lambda_V)$
- n_C and n_V are independent

where N_C is the extent of exposure for controls and N_V is the extent of exposure for vaccine.

n_v given $(n_v + n_c) = n$ as a conditional distribution is

$$\text{Bin}\left(n = n_v + n_c, P = \frac{\lambda_v N_v}{\lambda_v N_v + \lambda_c N_c}\right)$$

$$\text{Then } P = \frac{\frac{\lambda_v}{\lambda_c}\left(\frac{N_v}{N_c}\right)}{\frac{\lambda_v}{\lambda_c}\left(\frac{N_v}{N_c}\right)+1} = \frac{RC}{RC+1} \text{ where } R = \frac{\lambda_v}{\lambda_c},\ C = \frac{N_v}{N_c}.$$

Use (n_v, n_c) or $p = n_v/(n_v + n_c)$ to produce a $100(1 - \alpha)\%$ confidence interval (P_L, P_U) for P.

Use $\frac{P}{(1-P)C}$ or $\frac{N_c P}{(1-P)N_v}$ as estimator for R and

$\left\{\frac{P_L}{(1\text{-}P_L)C}, \frac{P_U}{(1\text{-}P_U)C}\right\}$ as $100(1-\alpha)\%$ confidence interval for $R = \lambda_v / \lambda_c$.

Example

Treatment	Events	Person Years
Vaccine	3	7500
Placebo	58	7250

```
data vaccine2;
   input Outcome $ Count @@;
datalines;
fail 3 success 58;
;

ods select BinomialCLs;
proc freq order=data;
  weight count;
  tables Outcome / binomial (exact);
  ods output BinomialCLs=BinomialCLs;
run;
```

The following IML code then produces the exact confidence interval for the IDR:

```
proc iml;
   use BinomialCLs var{lowerCL upperCL};
   read all into CL;
   print CL;
   q = { 3, 7500, 58, 7250 };
   C = q[2]/q[4];
   P = q[1]/(q[1]+q[3]);
   R = P/((1-P)*C);
   CI = CL[1]/((1-CL[1])*C) || CL[2]/((1-CL[2])*C);
   print r CI;
quit;
```

- The IDR that compares vaccine failure to placebo is 0.0500, with an exact confidence interval of (0.01002, 0.15355). The vaccine is much more effective than the placebo.

- The percent rate reduction in failures is computed as:

$$100(1 - \text{IDR}) = 95\%$$

with (84.645%, 98.998%) as the exact 95% confidence interval .

Exact Method for the Confidence Interval for the Odds Ratio for Association between a Dichotomous Factor (i.e., Two Groups) and a Dichotomous Response (i.e., Two Outcomes)

Let $i = 1,2$ index two groups from which independent simple random samples are selected.

Let y_i denote the number of subjects with favorable outcome in a sample of size n_i from group i.

Let π_i denote the fraction of subjects with favorable outcome in the population for the ith group (and thereby the probability that a randomly selected subject from group i has the favorable outcome).

The likelihood for (y_1, y_2) is the product of two binomial likelihoods.

$$\Pr\left\{(y_1, y_2)\middle|(\pi_1, \pi_2)\right\} = \frac{n_1!}{y_1!(n_1 - y_1)!}\pi_1^{y_1}(1-\pi_1)^{n_1-y_1}\frac{n_2!}{y_2!(n_2 - y_2)!}\pi_2^{y_2}(1-\pi_2)^{n_2-y_2}$$

Describe the variation between π_1 and π_2 with a logistic regression model:

$$\pi_1 = \exp(\alpha + \beta)/\{1 + \exp(\alpha + \beta)\}$$

$$\pi_2 = \exp(\alpha)/\{1 + \exp(\alpha)\}$$

Note that $\pi_1(1-\pi_2)/(1-\pi_1)\pi_2 = \exp(\beta) = \psi$

(the odds ratio)

Express the likelihood for (y_1, y_2) in terms of α and β from the logistic regression model:

$$\Pr\left\{\left(y_1, y_2\right)\middle|\left(\alpha, \beta\right)\right\} = \frac{n_1!}{y_1!(n_1-y_1)!}\frac{n_2!}{y_2!(n_2-y_2)!}\frac{\exp\left[(\alpha+\beta)y_1\right]\exp(\alpha y_2)}{\left[1+\exp(\alpha+\beta)\right]^{n_1}\left[1+\exp(\alpha)\right]^{n_2}}$$

The likelihood for $y_+ = \left(y_1 + y_2\right)$ is:

$$\Pr\left\{\left(y_+ \middle|\pi_1, \pi_2\right)\right\} = \sum_{g_1+g_2=y_+}\frac{n_1!}{g_1!(n_1-g_1)!}\frac{n_2!}{g_2!(n_2-g_2)!}\frac{\exp\left[y_+\alpha\right]\exp\left[g_1\beta\right]}{\left[1+\exp(\alpha+\beta)\right]^{n_1}\left[1+\exp(\alpha)\right]^{n_2}}$$

The conditional likelihood for (y_1, y_2) given y_+ and β is:

$$\Pr\left\{ y_1 \middle| y_+ \text{ and } \beta \right\} = \frac{\frac{n_1!}{y_1!(n_1-y_1)!}\frac{n_2!}{(y_+-y_1)!(n_2-y_++y_1)!}\exp(\beta y_1)}{\sum_{g_1=\max(0,y_+-n_2)}^{\min(n_1,y_+)} \exp(\beta g_1)\frac{n_1!}{g_1!(n_1-g_1)!}\frac{n_2!}{(y_+-g_1)!(n_2-y_++g_1)!}}$$

$$= \frac{\frac{n_1!}{y_1!(n_1-y_1)!}\frac{n_2!}{(y_+-y_1)!(n_2-y_++y_1)!}\exp(\beta y_1)}{\sum_{g_1=\max(0,y_+-n_2)}^{\min(n_1,y_+)} \left[\frac{n_1!}{g_1!(n_1-g_1)!}\frac{n_2!}{(y_+-g_1)!(n_2-y_++g_1)!}\exp(\beta g_1)\right]}$$

$= f\left(y_1 \middle| n_1, n_2, y_+, \beta\right) \quad = f\left(y_1 \middle| \beta\right)$ where $\beta = \log_e \psi$

for the non-central hypergeometric distribution

$$\sum_{g=\max(0, y_+ - n_2)}^{y_1} f\left(g \middle| \beta\right) = \Pr\left[Y_1 \le y_1 \middle| \beta\right]$$

which decreases as β increases, and so the largest β that this quantity is $\ge \alpha/2$ is the upper confidence limit β_U for β;

also, $$\sum_{g=y_1}^{\min(n_1, y_+)} f\left(g \middle| \beta\right) = \Pr\left[Y_1 \ge y_1 \middle| \beta\right]$$

decreases as β decreases, and so the smallest β that this quantity is $\ge \alpha/2$ is the lower confidence limit β_L for β.

$\exp(\beta_L, \beta_U) = (\psi_L, \psi_U)$ is the corresponding confidence interval for the odds ratio.

When $\beta = 0$ (and so $\pi_1 = \pi_2 = \pi_*$),

$$\Pr\left\{ y_+ \,\middle|\, \pi_1 = \pi_2 \right\} = \frac{(n_1 + n_2)!}{y_+!(n_1 + n_2 - y_+)!} \pi_*^{y_+} (1 - \pi_*)^{n_+ - y_+} =$$

$$\frac{\exp(y_+ \alpha)}{\left[1 + \exp(\alpha)\right]^{n_+}} \left\{ \sum_{g_1 = \max(0, y_+ - n_2)}^{\min(n_1, y_+)} \frac{n_1!}{g_1!(n_1 - g_1)} \frac{n_2!}{(y_+ - g_1)!(n_2 - y_+ + g_1)!} \right\}$$

where $n_+ = n_1 + n_2$. But correspondingly

$$\left\{ \sum_{g_1 = \max(0, y_+ - n_+ + n_1)}^{\min(n_1, y_+)} \frac{n_1!}{g_1!(n_1 - g_1)!} \frac{(n_+ - n_1)!}{(y_+ - g_1)!(n_+ - n_1 - y_+ + g_1)!} \right\}$$

$$= \frac{(n_+)!}{y_+!(n_+ - y_+)!}$$

and so when $\beta = 0$,

$$\Pr\left\{y_1 \middle| n_+, n_1, y_+\right\} = \frac{\left\{\frac{n_1!}{y_1!(n_1-y_1)!}\frac{(n_+-n_1)!}{(n_+-n_1-y_++y_1)!(y_+-y_1)!}\right\}}{\left\{\frac{n_+!}{y_+!(n_+-y_+)!}\right\}}$$

$$= \frac{n_1!(n_+-n_1)!y_+!(n_+-y_+)!}{n_+!y_1!(n_1-y_1)!(y_+-y_1)!(n_+-n_1-y_++y_1)!}$$

which is the hypergeometric distribution for Fisher's exact test.

Methodology for Exact Confidence Interval for Risk Difference (SAS 9.2)

An exact 100(1−α)% confidence interval for the risk difference can be obtained in SAS 9.2 by specifying an EXACT RISKDIFF statement in PROC FREQ.

Denote the proportion difference by $d = p_1 - p_2$. For a 2×2 table with row totals n_1 and n_2, the joint probability function (product of two independent binomials) can be expressed in terms of the table cell frequencies and the parameters d and p_2 as follows:

$$f(n_{11}, n_{21}; n_1, n_2, d, p_2) = \binom{n_1}{n_{11}} (d + p_2)^{n_{11}} (1 - d - p_2)^{n_1 - n_{11}} \times \binom{n_2}{n_{21}} (p_2)^{n_{21}} (1 - p_2)^{n_2 - n_{21}}$$

When constructing confidence limits for the proportion difference, the parameter of interest is d and p_2 is a nuisance parameter.

Denote the observed value of the proportion difference by $d_0 = \hat{p}_1 - \hat{p}_2$

The 100(1−α)% confidence limits for d (denoted d_L and d_U) are computed as

$$d_L = \sup(d_* : P_U(d_*) > \alpha/2)$$

$$d_U = \inf(d_* : P_L(d_*) > \alpha/2)$$

Where

$$P_U(d_*) = \sup_{p_2}\left(\sum_{A, D(a)\geq d_0} f(n_{11}, n_{21}; n_1, n_2, d_*, p_2)\right)$$

$$P_L(d_*) = \sup_{p_2}\left(\sum_{A, D(a)\leq d_0} f(n_{11}, n_{21}; n_1, n_2, d_*, p_2)\right)$$

The set A includes all 2×2 tables with row sums equal to n_1 and n_2, and D(a) denotes the value of the proportion difference ($p_1 - p_2$) for table a in A. To compute $P_U(d_*)$, the sum includes probabilities of those tables for which (D(a)≥d_0), where d_0 is the observed value of the proportion difference. For a fixed value of d_*, $P_U(d_*)$ is taken to be the maximum sum over all possible values of p_2. Details can be found in Santner and Snell (1980) and Agresti and Min (2001).

Essentially, this method determines the greatest lower confidence limit and smallest upper confidence limit such that, with row totals fixed, the sum of table probabilities where risk differences are as extreme or more extreme than the observed risk difference are no more than $\alpha/2$ in either direction.

The confidence limits are conservative for small samples because this is a discrete problem; the confidence coefficient is not exactly $1-\alpha$ but is at least $1-\alpha$ (Agresti, 1992).

SAS Version 9

PROC FREQ: TABLES Statement Options

Option	***Description***
CHISQ	chi-square tests and measures of association based on chi-square
MEASURES	measures of association and their asymptotic standard errors
AGREE	tests and measures of classification agreement, including McNemar's for 2×2 tables and kappa statistics
ALL	requests tests and measures of association produced by CHISQ, MEASURES, and CMH
FISHER	Fisher's exact test for tables larger than 2×2

CL	confidence limits for the MEASURES statistics
EXPECTED	displays the expected cell frequency for each cell
NOROW	suppresses display of the row percentage for each cell
NOCOL	suppresses display of the column percentage for each cell
NOPERCENT	suppresses display of the overall percentage for each cell
ALPHA=	sets the confidence level for confidence limits

SAS Version 9.3

PROC FREQ: TABLES Statement Options for Difference in Proportions

Option	***Description***
RISKDIFF	requests probabilities, differences in probabilities, and their CI's for 2×2 tables
RISKDIFF (CORRECT)	RISKDIFF statistics with a continuity correction for all statistics for which it is applicable
RISKDIFF (CL=(WALD) CORRECT)	Include corrected Wald confidence intervals
RISKDIFF (CL=(MN))	Include Miettinen-Nurminem confidence interval (Sometimes preferred when cell count size is marginal)
RISKDIFF (CL=(NEWCOMBE))	Include Newcombe confidence interval (Sometimes preferred for small cell counts)
RISKDIFF (CL=(EXACT))	Include unconditional exact confidence interval

SAS Version 9.2
PROC FREQ: EXACT Statement Options

Option	***Exact Statistics Computed***
CHISQ	Pearson chi-square, likelihood-ratio chi-square, and Mantel-Haenszel chi-square tests
RISKDIFF	confidence limits for risk difference in 2×2 tables
MEASURES	tests for the Pearson correlation and the Spearman correlation, and the odds ratio confidence limits for 2×2 tables
OR	confidence limits for the odds ratio for 2×2 tables
FISHER	Fisher's exact test
BINOMIAL	confidence limits for a single binomial proportion
MCNEM	McNemar's test
AGREE	McNemar's test for 2×2 tables, simple kappa coefficient, and weighted kappa coefficient

9. Sample Size

Reference, Fleiss [1981]

Based on Counterpart to Pearson χ^2 Statistic (2-Sample)

Fisher's exact test results are better approximated if sample size is based on continuity-corrected counterpart to the Pearson chi-square statistic.

1. For equal sample size in both groups:

$$n = \frac{n'}{4}\left[1+\sqrt{1+\frac{4}{n'(\pi_1-\pi_2)}}\right]^2 \approx n' + \frac{2}{|\pi_1-\pi_2|}$$

where

$$n' = \frac{\left[z_\alpha\sqrt{2\bar{\pi}(1-\bar{\pi})}+z_\beta\sqrt{\pi_1(1-\pi_1)+\pi_2(1-\pi_2)}\right]^2}{(\pi_1-\pi_2)^2}$$

for which $\bar{\pi} = \frac{\pi_1+\pi_2}{2}$, and z_α and z_β are the $100(1-\alpha)$th and $100(1-\beta)$th percentiles of $N(0,1)$

2. For a specified sample size of n per group, power to test $H_0 : \pi_1 = \pi_2$ against $H_1 : (\pi_1 - \pi_2) = \Delta > 0$ is determined via

$$n' = n - \frac{2}{|\pi_1 - \pi_2|}$$

$$z_\beta = \frac{|\pi_1 - \pi_2|\sqrt{n'} - z_\alpha \sqrt{2\bar{\pi}(1-\bar{\pi})}}{\sqrt{\pi_1(1-\pi_1) + \pi_2(1-\pi_2)}}$$

$$\text{So, Power} = \Pr\left\{\text{Std. Normal} \le \frac{|\pi_1 - \pi_2|\sqrt{n'} - z_\alpha \sqrt{2\bar{\pi}(1-\bar{\pi})}}{\sqrt{\pi_1(1-\pi_1) + \pi_2(1-\pi_2)}}\right\}$$

3. For two-sided tests, apply methods here with $\frac{\alpha}{2}$

4. For the general situation with n_1 subjects in Group 1 and $n_2 = kn_1$ in Group 2

$$n_1' = \frac{\left[z_\alpha \sqrt{(k+1)\bar{\pi}(1-\bar{\pi})} + z_\beta \sqrt{k\pi_1(1-\pi_1) + \pi_2(1-\pi_2)} \right]^2}{k(\pi_1 - \pi_2)^2}$$

$$n_1 = \frac{n_1'}{4}\left[1 + \sqrt{1 + \frac{2(k+1)}{n_1' k \left| \pi_1 - \pi_2 \right|}} \right]^2$$

$$\approx n_1' + \frac{(k+1)}{k\left| \pi_1 - \pi_2 \right|},$$

where $\bar{\pi} = \frac{(\pi_1 + k\pi_2)}{(1+k)}$

5. For the situation with n subjects in one group which corresponds to π_1 , the sample size to test π_2 as a null hypothesis is given by #4 above with $k = \infty$, i.e.:

$$n_1' = \frac{\left[z_\alpha \sqrt{\pi_2(1-\pi_2)} + z_\beta \sqrt{\pi_1(1-\pi_1)} \right]^2}{(\pi_1 - \pi_2)^2}$$

$$n_1 = \frac{n_1'}{4}\left[1 + \sqrt{1 + \frac{2}{n'|\pi_1 - \pi_2|}} \right]^2$$

Example for sample size calculations: Suppose a clinical trial is being planned to compare a test treatment, an active control treatment, and placebo for the healing of ulcers. Suppose the expected healing rates are $\pi_1 = 0.45$ for placebo, $\pi_2 = 0.68$ for the active control, and $\pi_3 = 0.82$ for the test treatment.

a) What sample size per treatment is needed to have 0.80 power for the comparison between the test treatment and the active control at the two-sided $\alpha = 0.05$ significance level (given that the sample sizes are equal for these two treatments)?

For the general situation with n_1 subjects in Group 1 and $n_2 = kn_1$ subjects in Group 2:

$$n_1' = \frac{\left[z_\alpha \sqrt{(k+1)\bar{\pi}(1-\bar{\pi})} + z_\beta \sqrt{k\pi_1(1-\pi_1) + \pi_2(1-\pi_2)} \right]^2}{k(\pi_1 - \pi_2)^2}$$

$$n_1 = \frac{n_1'}{4}\left[1 + \sqrt{1 + \frac{2(k+1)}{n_1' k \left|\pi_1 - \pi_2\right|}} \right]^2 \approx n_1' + \frac{(k+1)}{k\left|(\pi_1 - \pi_2)\right|}$$

where $\bar{\pi} = \dfrac{(\pi_1 + k\pi_2)}{(1+k)}$

For our problem, $k = 1$, so $\bar{\pi} = \dfrac{(\pi_2 + \pi_3)}{2} = \dfrac{(0.68 + 0.82)}{2} = 0.75$

$$n' = \frac{\left[1.96\sqrt{2(0.75)(0.25)} + 0.84\sqrt{0.68(0.32) + 0.82(0.18)}\right]^2}{(0.68 - 0.82)^2} = 148.8$$

and

$$n = \frac{148.8}{4}\left[1 + \sqrt{1 + \frac{4}{148.8|0.68 - 0.82|}}\right]^2$$

$$= 162.8 \approx 163 \text{ per group}$$

b) What sample size is needed to provide 0.80 power for the comparison between the active control and placebo at the two-sided $\alpha = 0.05$ significance level in a research design where twice as many patients receive active control as placebo?

n_1 patients in the placebo arm

$\Rightarrow k = 2$

$n_2 = 2n_1$ patients in the active drug arm

$$\bar{\pi} = \frac{(\pi_1 + k\pi_2)}{(k+1)} = \frac{(0.45 + (2)0.68)}{3} = 0.60$$

$$n_1' = \frac{\left[1.96\sqrt{3(0.60)(0.40)} + 0.84\sqrt{2(0.45)(0.55) + 0.68(0.32)}\right]^2}{2(0.45-0.68)^2}$$

$$= 53.2$$

and

$$n_1 = \frac{53.2}{4}\left[1 + \sqrt{1 + \frac{6}{2(53.2)\left|0.45-0.68\right|}}\right]^2$$

= 59.5 ≈ 60 patients in placebo group and
120 patients in active group

c) Suppose the rate of an adverse event is 1% in the treatment group. How many subjects are needed to have 90% power at the one-sided $\alpha = 0.025$ level so that the rate of the adverse event is less than 3%?

$$n' = \frac{\left[1.96\sqrt{(.03)(1-.03)} + 1.282\sqrt{(.01)(1-.01)}\right]^2}{(.01-.03)^2} = 534$$

$$n = \frac{534}{4}\left[1 + \sqrt{1 + \frac{2}{(534)|0.01-0.03|}}\right]^2 = 583$$

Examples of sample size calculations using PROC POWER

a) By default, PROC POWER conducts a two-sided test with an 0.05 significance level. Setting NTotal equal to missing tells SAS to solve for the overall sample size.

```
proc power;
    twosamplefreq test=pchi
    grouppproportions= (.68 .82)
    power=.8
    ntotal=.;
run;
```

The POWER Procedure

Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements

Distribution	Asymptotic normal
Method	Normal approximation
Alpha	0.05
Group 1 Proportion	0.68
Group 2 Proportion	0.82
Group 1 Weight	1
Group 2 Weight	1
Nominal Power	0.8
Number of Sides	2
Null Proportion Difference	0

Computed N Total

Actual Power	N Total
0.800	298

- Total *N* is 298, so *n* = 149 per group.

- The statement TEST=PCHI tells SAS to use Pearson chi-square which is computed without continuity correction

- Invoking TEST=FISHER tells SAS to use Fisher's test, which produces a result more similar to the computation for Pearson's chi-square test with continuity correction

```
proc power;
   twosamplefreq test=fisher
   grouppproportions= (.68 .82)
   power=.8
   npergroup=.;
run;
```

The POWER Procedure

Fisher's Exact Conditional Test for Two Proportions

Fixed Scenario Elements

Distribution	Exact conditional
Method	Walters normal approximation
Group 1 Proportion	0.68
Group 2 Proportion	0.82
Nominal Power	0.8
Number of Sides	2
Alpha	0.05

Computed N Per Group

Actual Power	N Per Group
0.801	162

- *n* per group is 162.

b) This question asks for unequal sample sizes, so the GROUPWEIGHTS option is added, indicating a two-to-one ratio for subjects in the active control group compared to the placebo group.

```
proc power;
  twosamplefreq test=pchi
  grouppproportions= (.45 .68)
  power=.8
  groupweights= (1 2)
  ntotal=.;
run;
```

The POWER Procedure

Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements

Distribution	Asymptotic normal
Method	Normal approximation
Alpha	0.05
Group 1 Proportion	0.45
Group 2 Proportion	0.68
Group 1 Weight	1
Group 2 Weight	2
Nominal Power	0.8
Number of Sides	2
Null Proportion Difference	0

Computed N Total

Actual Power	N Total
0.806	162

- Overall sample size is $N = 162$. So $n = 54$ in placebo group and $n = 108$ in active control group.

- NOTE: nQuery could also be used for these sample size calculations, specifying Pearson's chi-square either with or without continuity correction

- Again, invoking TEST=FISHER tells SAS to use Fisher's test, which produces a result more similar to the computation for Pearson's chi-square test with continuity correction

```
proc power;
      twosamplefreq test=fisher
      grouppproportions= (.45 .68)
      power=.8
      groupweights= (1 2)
      ntotal=.;
run;
```

The POWER Procedure

Fisher's Exact Conditional Test for Two Proportions

Fixed Scenario Elements

Distribution	Exact conditional
Method	Walters normal approximation
Group 1 Proportion	0.45
Group 2 Proportion	0.68
Group 1 Weight	1
Group 2 Weight	2
Nominal Power	0.8
Number of Sides	2
Alpha	0.05

Computed N Total

Actual Power	N Total
0.807	183

- Overall sample size is $N = 183$,
- $n = 61$ in placebo group; $n = 122$ in active control group

c) Using $n = 75$ for the placebo group and $n = 149$ for test treatment, specify the different n's using the GROUPNS option. To solve for the power, set the POWER option equal to missing.

```
proc power;
     twosamplefreq test=pchi
     grouppproportions = (.45 .82)
     power=.
     groupns = (75 149);
run;
```

The POWER Procedure

Pearson Chi-square Test for Two Proportions

Fixed Scenario Elements

Distribution	Asymptotic normal
Method	Normal approximation
Group 1 Proportion	0.45
Group 2 Proportion	0.82
Group 1 Sample Size	75
Group 2 Sample Size	149
Number of Sides	2
Null Proportion Difference	0
Alpha	0.05

Computed Power

Power
>.999

- So the power would be > 99%

- NOTE: nQuery could also be used for this power calculation, using Fisher's Exact Test or Pearson's chi-square with or without continuity correction.

- TEST=FISHER

```
proc power;
     twosamplefreq test=fisher
     grouppproportions = (.45 .82)
     power=.
     groupns = (75 149);
run;
```

The POWER Procedure

Fisher's Exact Conditional Test for Two Proportions

Fixed Scenario Elements

Distribution	Exact conditional
Method	Walters normal approximation
Group 1 Proportion	0.45
Group 2 Proportion	0.82
Group 1 Sample Size	75
Group 2 Sample Size	149
Number of Sides	2
Alpha	0.05

Computed Power

Power
>.999

d) This question is only dealing with one group, so we use the ONESAMPLEFREQ statement.

- Using TEST=ADJZ specifies a normal-approximate *z*-test with continuity adjustments, and METHOD=NORMAL computes approximate results using the normal approximation to the binomial distribution

- SIDES = 1 is used to indicate a one-sided test with the alternative hypothesis in the same direction as the effect; you could also use SIDES=L to test that the alternative is less than the null value

```
proc power;
        onesamplefreq test=adjz method=normal
        alpha= .025
        sides=1
        nullproportion=.03
        proportion=.01
        power=.90
        ntotal=.;
run;
```

The POWER Procedure
Z Test for Binomial Proportion with Continuity Adjustment

Fixed Scenario Elements

Method	Normal approximation
Number of Sides	1
Null Proportion	0.03
Alpha	0.025
Binomial Proportion	0.01
Nominal Power	0.9

Computed N Total

Actual Power	N Total
0.901	583

- So the total sample size needed is *N*=583 subjects.