PharmaSUG China 3rd Conference, 2014

Calculation of Confidence Intervals for Relative Risk and Odds Ratio Using StatXact PROCs

Halimu Haridona, PPD Inc., Beijing

ABSTRACT

In clinical trials, we will face some issues when making statistical inference using asymptotic and approximate statistical methods for rare events, e.g., calculating the confidence intervals of relative risk/odds ratio for low incidence binomial endpoints. This paper will introduce how to calculate the relative risk and odds ratio and their confidence intervals using StatXact PROCs, which is a more reliable method. StatXact PROCs is a statistical package for Exact Nonparametric Inference, developed for SAS users. It is used to make reliable inferences by exact and Monte Carlo methods when the data are sparse, heavily tied, or skewed, and the accuracy of the corresponding large sample theory is in doubt.

INTRODUCTION

SAS StatXact PROCs is a statistical package for Exact Nonparametric Inference. The goal of StatXact PROCs is to enable statisticians and data analysts to make reliable inferences by exact and Monte Carlo methods when their data are sparse, heavily tied, or skewed, and the accuracy of the corresponding large sample theory is in doubt.

When making the statistical reference, some basic assumptions will be made on the underlying distribution of the data, e.g., the normal distribution for continuous data, binomial, Poisson, or chi-square distribution for categorical data. Nevertheless, for both the continuous and categorical cases, the assumption is sometimes difficult to verify. They assume that the data set is large enough for the test statistic to converge to an appropriate limiting normal or chi-square distribution. P-values are then obtained by evaluating the tail area of the limiting distribution, instead of actually deriving the true distribution of the test statistic and then evaluating its tail area. P-values based on the large-sample assumption are known as asymptotic p-values, while p-values based on deriving the true distribution of the test statistic are termed exact p-values. While one would prefer to use exact p-values for scientific inference they often pose formidable computational problems and so, as a practical matter, asymptotic p-values are used in their place. For large and well-balanced data sets this makes very little difference since the exact and asymptotic p-values are very similar. But for small, sparse, unbalanced, and heavily tied data, the exact and asymptotic p-values can be quite different and may lead to opposite conclusions concerning the hypothesis of interest.

There are two basic types of algorithms in StatXact PROCs, complete enumeration and Monte Carlo enumeration. The complete enumeration algorithms enumerate every single outcome in the reference set. Thus they always produce the exact p-value. Their result is 100% accurate. They are not, however, guaranteed to solve every problem. Some data sets might be too large for complete enumeration of the reference set. For this reason we also provide Monte Carlo enumeration algorithms. These algorithms enumerate a random subset of all the possible outcomes in the reference set. The Monte Carlo algorithms provide an estimate of the exact p-value, called the Monte Carlo p-value, which can be made as accurate as necessary for the problem on hand. Typically they provide a 99% accurate answer, but the user is free to set the level of accuracy to any arbitrary degree. For a given level of accuracy the range within which the p-value is located can be narrowed simply by sampling more outcomes from the reference set. Moreover Monte Carlo methods are guaranteed to solve any problem, no matter how large the data set. Thus they provide a robust, reliable back-up for the situations in which the complete enumeration alogrithms fail.

Finally the asymptotic p-value is always available by default in StatXact PROCs. This computation is straightforward, and does not require any special algorithms.

EXACT CONFIDENCE INTERVAL FOR THE ODDS RATIO

We may be interested in comparing the occurrence of some endpoints with very low incidence rate between two treatment groups when analyzing the clinical trial data, e.g., comparing the incidence rate of local steroid effect in asthma studies. The odds ratio and its confidence interval can't be calculated by commonly used SAS FREQ procedure if no event happened (zero count) in either treatment group. But this effort can be handled by StatXact PROCs very efficiently.

Below is a dummy data, variable TRTPN is for treatment group (1 = study treatment group, 2 = control group), variable OCCUR identifies whether the event occurred (1 = yes, 0 = no), variable COUNT is count of subjects.

The SAS codes for input dataset, frequency procedure, and StatXact binomial procedure are as below. The odds/ex option in binomial procedure is for exact odds ratio, po is for population variable (treatment group variable), ou is for outcome variable.

```
data dummy1;
    input TRTPN $ OCCUR $ COUNT;
    cards;
    1 1 0
    1 0 300
    2 1 9
    2 0 300
    ;
run;
proc freq data=dummy1;
    tables TRTPN*OCCUR/relrisk cmh;
    weight COUNT;
run;
proc binomial data=dummy1;
    odds/ex;
    po TRTPN;
    ou OCCUR;
    weight COUNT;
run;
```

Since we have zero count for study treatment group so the odds ratio and the 95% confidence interval can't be calculated in frequency procedure, see outputs as below. The Cochran-Mantel-Haenszel statistics are requested in the frequency procedure since this is the corresponding asymptotic method for the exact P value when calculating the odds ratio using binomial procedure.

The FREQ Procedure

Summary Statistics for TRTPN by OCCUR

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

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	8.8544	0.0029
2	Row Mean Scores Differ	1	8.8544	0.0029
3	General Association	1	8.8544	0.0029

Estimates of the Common Relative Risk (Row1/Row2)

Type of Study	Method	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	Mantel-Haenszel	.	.	.
	Logit **	19.0000	1.1009	327.9120
Cohort (Col1 Risk)	Mantel-Haenszel	1.0300	1.0103	1.0501
	Logit	1.0300	1.0103	1.0501
Cohort (Col2 Risk)	Mantel-Haenszel	0.0000	.	.
	Logit **	0.0542	0.0032	0.9272

To avoid undefined results, some estimates are not computed.
** These logit estimators use a correction of 0.5 in every cell of those tables that contain a zero.

Total Sample Size = 609

However, the odds ratio (control group/study treatment group) and 95% CI are calculated by StatXact BINOMIAL procedure. The exact odds ratio (95% CI) is 0.0000 (0.0000, 0.5129), exact P value is 0.0042 (showing significant difference between treatment groups). The asymptotic P value is provided (same as CMH P value from frequency procedure) but the asymptotic 95% CI for odds ratio is undefined. The exact mid-p adjusted P value and 95% CI for

odds ratio are also provided. Mid-p correction is a conservative adjustment method which provides shorter confidence interval. In general, StatXact recommends the use of the mid-p confidence intervals.

```
Output from StatXact (r) PROCs (v10.0) _SAS9_2
Copyright (c) 1987-2013 Cytel Inc., Cambridge, MA, USA.

ODDS RATIO OF TWO BINOMIAL PROPORTIONS

Data file name : < DUMMY >
Population Variable Name : TRTPN
Outcome Variable Name : OCCUR
Weight Variable Name : COUNT


Statistic based on the observed  2 by  2 table :

    Binomial proportion for column <        1> : pi_1   =      1.0000
    Binomial proportion for column <        2> : pi_2   =      0.9709

                               ( pi_2 )/( 1-pi_2 )
    Odds Ratio                = ------------------   =      0.0000
                               ( pi_1 )/( 1-pi_1 )
Results:
-------------------------------------------------------------------------
Method                    P-value(2-sided)       95.00 Confidence Interval
-------------------------------------------------------------------------
Asymp (Mantel-Haenszel)         0.0029        (   Undefined ,   Undefined)
Exact                           0.0042        (      0.0000 ,      0.5129)
Exact-Mid P                     0.0032        (      0.0000 ,      0.3993)
```

As we discussed in the introduction section, the results of exact and asymptotic method can be quite different for small, sparse, unbalanced, and heavily tied data, i.e., asymptotic method is unreliable for those data. In above example, if we have no zero count in both treatment groups, the odds ratio and confidence interval will be calculated by asymptotic method, but it is not a good estimate of exact (true) results. Let's see another example.

```
data dummy2;
    input TRTPN $ OCCUR $ COUNT;
    cards;
    1 1 1
    1 0 100
    2 1 8
    2 0 100
    ;
run;

proc binomial data=dummy2;
    odds/ex;
    po TRTPN;
    ou OCCUR;
    weight COUNT;
run;
```

```
Output from StatXact (r) PROCs (v10.0) _SAS9_2
Copyright (c) 1987-2013 Cytel Inc., Cambridge, MA, USA.

ODDS RATIO OF TWO BINOMIAL PROPORTIONS

Data file name : < DUMMY2 >
Population Variable Name : TRTPN
Outcome Variable Name : OCCUR
Weight Variable Name : COUNT


Statistic based on the observed  2 by  2 table :

    Binomial proportion for column <        1> : pi_1   =      0.9901
    Binomial proportion for column <        2> : pi_2   =      0.9259

                               ( pi_2 )/( 1-pi_2 )
    Odds Ratio                = --------------------   =      0.1250
                               ( pi_1 )/( 1-pi_1 )
Results:
-------------------------------------------------------------------------
Method                    P-value(2-sided)       95.00 Confidence Interval
-------------------------------------------------------------------------
Asymp (Mantel-Haenszel)         0.0227        (      0.0153 ,      1.0180)
Exact                           0.0448        (      0.0028 ,      0.9685)
Exact-Mid P                     0.0347        (      0.0055 ,      0.8078)
```

The odds ratio (control group/study treatment group) is 0.1250, the asymptotic 95% CI (0.0153, 1.0180) is quite different from exact 95% CI (0.0028, 0.9685), statistical inferences from these two methods are opposite. Therefore, if we rely on the asymptotic method, we will draw wrong conclusion.

EXACT CONFIDENCE INTERVAL FOR THE RELATIVE RISK

Exact Test of Non-inferiority

Chan (1998) discusses a vaccine efficacy study of a recombinant DNA Influenza A vaccine against wild-type H1N1 virus challenge. The study compares the infection rates in the vaccinated and placebo groups. There were 15 individuals in each group. The following data (in Table 1) were obtained.

Table 1

	Treatment Group		
Disease Status	Placebo	Vaccine	Total
Infected	12 (80%)	7 (47%)	19
Not Infected	3 (20%)	8 (53%)	11
Total	15	15	30

Let $\pi 2$ be the infection rate in the vaccinated group and $\pi 1$ be the infection rate in the placebo group. Define $\rho = \pi 2/\pi 1$, and define $\lambda = 1 - \rho$. The parameter λ is known as the vaccine efficacy. We assume that $\pi 2 \leqslant \pi 1$. Therefore the new vaccine has 100% efficacy if $\pi 2 = 0$ and no efficacy if $\pi 2 = \pi 1$. Barnard's test of superiority was a test of the null hypothesis that $\lambda = 0$. That hypothesis was rejected in favor of the alternative that $\lambda > 0$. From a public health standpoint, however, the benefits from vaccination must exceed a given threshold in order to justify the risk of vaccinating healthy subjects. Therefore, in designing vaccine trials, one typically chooses a non-zero efficacy lower bound. Suppose we choose $\lambda_0 = 0.1$ as the non-zero efficacy lower bound. This implies that if $\lambda \leqslant 0.1$, the virus does not offer sufficient benefit relative to placebo to justify using it on a large scale for the prevention of infection. Thus we wish to test the null hypothesis of insufficient vaccine efficacy (i.e., inferiority) $\lambda \leqslant 0.1$ against the one-sided alternative hypothesis of sufficient vaccine efficacy (i.e., non-inferiority), $\lambda > 0.1$. Equivalently, we wish to test the null hypothesis of inferiority,

H0: $\rho \geqslant 0.9$ versus ,

against the alternative hypothesis of non-inferiority

H1: $\rho < 0.9$.

Now Execute the Non-Inferiority Test by using the following SAS code.

```
data vaccine;
    input TREAT $ DISEASE $ COUNT @@;
    cards;
    PLACEBO INF 12
    PLACEBO NOINF 3
    VACCINE INF 7
    VACCINE NOINF 8
run;
proc binomial data=vaccine;
    noninf/ex ratio margin=0.9;
    po TREAT;
    ou DISEASE;
    weight COUNT;
run;
```

The output of StatXact is as below:

```
Output from StatXact (r) PROCs (v10.0) _SAS9_2
Copyright (c) 1987-2013 Cytel Inc., Cambridge, MA, USA.

UNCONDITIONAL TEST OF NON-INFERIORITY USING RATIO OF TWO BINOMIAL PROPORTIONS

Data file name : < VACCINE >
Population Variable Name : TREAT
Outcome Variable Name : DISEASE
Weight Variable Name : COUNT

    H0:(pi_2/pi_1) .GE. rho_0 vs H1: (pi_2/pi_1) .LT. rho_0

Statistic based on the observed  2 by  2 table :

    Observed proportion for population <        1> : piHat_1                  =      0.8000
    Observed proportion for population <        2> : piHat_2                  =      0.4667
    Observed ratio of proportions : piHat_2/piHat_1                           =      0.5833
    Maximum  margin of non-inferiority  : pi_2/pi_1 = rho_0                   =      0.9000
    Stderr(restricted mle of stdev of piHat_2-piHat_1*rho_0 given rho_0)      =      0.1689
    Standardized test statistic: (piHat_2-piHat_1*rho_0)/Stderr               =     -1.4999

Results:
-----------------------------------------------------------------------
               1-sided P-value              97.50% Upper Confidence
     Method      Pr{T .LE. t}                Bound for pi_2/pi_1
-----------------------------------------------------------------------
      Asymp            0.0668                      1.0294
      Exact            0.0857                      1.0372
```

The exact one-sided non-inferiority p-value is 0.0857, implying that the null hypothesis of insufficient vaccine efficacy cannot be rejected at the 2.5% level of significance. (Note: It is a common practice to adopt the 2.5% significance level for a 1-sided test and the 5% significance level for a two-sided test). The corresponding 97.5% upper confidence bound for ρ is 1.037 and confirms that we cannot rule out the possibility that $\rho > 0.9$.

Exact Test of Equivalence

Consider the following hypothetical dataset (Table 2):

Table 2

Disease Status	Treatment Group		Total
	Drug-1	Drug-2	
Response	94 (32%)	100 (33%)	194
No Response	200 (68%)	200 (67%)	400
Total	294	300	594

Let $\rho = \max(\pi 1/\pi 2, \pi 2/\pi 1)$. Suppose that Drug-1 and Drug-2 are regarded as equivalent if $\rho < 1.3$. The above data can be analyzed by using the following SAS code.

```
data drug12;
     input TREAT $ STATUS $ COUNT @@;
     cards;
     DRUG1 RESP-1 94
     DRUG1 RESP-2 200
     DRUG2 RESP-1 100
     DRUG2 RESP-2 200
run;
proc binomial data=drug12 ti=15 ;
     eqv/ex ratio margin=1.3;
     po TREAT;
     ou STATUS;
     weight COUNT;
run;
```

The output of StatXact is as below:

```
Output from StatXact (r) PROCs (v10.0) _SAS9_2
Copyright (c) 1987-2013 Cytel Inc., Cambridge, MA, USA.

UNCONDITIONAL TEST FOR EQUIVALENCE USING RATIO OF TWO BINOMIAL PROPORTIONS

Data file name : < DRUG12 >
Population Variable Name : TREAT
Outcome Variable Name : STATUS
Weight Variable Name : COUNT

H01:(pi_1/pi_2) .GE. rho_0 vs H11: (pi_1/pi_2) .LT. rho_0
H02:(pi_2/pi_1) .GE. rho_0 vs H12: (pi_2/pi_1) .LT. rho_0

Statistics based on the observed 2 by 2 table :

    Observed proportion for population <        1> : piHat_1                    =      0.3197
    Observed proportion for population <        2> : piHat_2                    =      0.3333
    Observed ratio of proportions : piHat_2/piHat_1                             =      1.0426
    Maximum  margin of equivalence : Max(pi_2/pi_1, pi_1/pi_2) = rho_0          =      1.3000
    Stderr(restricted mle of stdev of piHat_2-piHat_1*rho_0 given rho_0)        =      0.0440
    Standardized test statistic (t_01): (piHat_1-piHat_2*rho_0)/Stderr          =     -1.8688
    Standardized test statistic (t_02): (piHat_2-piHat_1*rho_0)/Stderr          =     -2.5842

Results:
-------------------------------------------------------------------------------
                1-sided P-value      1-sided P-value         95.00% Conf. Interval
    Method      Pr{T .LE. t_01}      Pr{T .LE. t_02}            for pi_2/pi_1
-------------------------------------------------------------------------------
     Asymp           0.0049               0.0308        (      0.8277,      1.3141)
     Exact           0.0051               0.0319        (      0.8268,      1.3159)
```

Based on the above output screen, we can reject H_{01} at the 2.5% level of significance, since the exact p-value for this hypothesis is 0.0051. We cannot, however, reject H_{02} at the 2.5% significance level since the exact p-value for this hypothesis is 0.0319. Thus the null hypothesis of inequivalence is not rejected at an overall 2.5% level of significance. This is also reflected in the 95% exact confidence interval for ρ (0.8268 to 1.3159), which contains the value of the equivalence margin, $\rho_0 = 1.3$.

CONCLUSION

Very efficient numerical algorithms are used in StatXact PROCs. For small data sets, this ensures quick computation of exact p-values, confidence intervals. If a data set is too large for exact inference, it is almost certainly large enough for the asymptotic theory to work accurately. In effect, StatXact PROCs provides a comprehensive set of nonparametric procedures that perform reliably for small, intermediate and large data sets.

REFERENCES

StatXact ® 10 PROCs for SAS® Users, Cytel Software Copyright 2013

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Halimu Haridona
Enterprise: PPD Inc.
Address: 8/F, Tower B, Central Point Plaza No 11, Dongzhimen South Ave. Dongcheng District
City, State ZIP: Beijing, 100007
Work Phone: +86 10-57636250
Fax: +86 10-57636251
E-mail: Halimu.Haridona@ppdi.com
Web: www.ppdi.com