

Computation of CIs for Binomial proportions in SAS and its practical difficulties

Jose Abraham, Kreara Solutions Pvt. Ltd., Thiruvananthapuram, India

ABSTRACT

In clinical trial data analysis, one of the most commonly used method for the analysis of categorical data is the Binomial proportion and their confidence intervals. If the variable of interest is a dichotomous categorical one, i.e. a variable with two levels (e.g. Subjects with serious AEs/Subjects without any serious AEs) then Binomial proportions (or percentages) and their CIs (or %CIs) can be used in the analysis. SAS[®] procedure which is usually used to obtain these proportions and CIs is the PROC FREQ. This paper describes the computation of Clopper Pearson CIs for binomial proportions using the PROC FREQ in SAS, practical difficulties in obtaining the correct CI values in situations like computing the CIs for responders when there are no responses available in the data, and also provides some programming tips to avoid erroneous CI computations.

INTRODUCTION

In the field of statistics, the Binomial distribution is commonly used in a variety of applications, especially in situations in which the outcomes of interest can be categorized into two, i.e. either a success or a failure. Binomial proportions and their confidence intervals are widely used in oncology trials to evaluate the response rate and are sometimes used in the other trials to assess the overall incidence of adverse events.

Confidence intervals for a binomial proportion can be estimated using various methods like Wald's (Normal Approximation) method or the Clopper-Pearson (exact) method.

NORMAL APPROXIMATION METHOD

The 'Normal approximation' method got its name because of the use of the z-value from the Normal Distribution. Normal approximation method is easy to compute and use of normal approximation method is supported by the central limit theorem and with sufficiently large sample size 'n', the Normal distribution is a good estimate of the Binomial distribution.

Equation to compute the Binomial CI using the Normal approximation method is given below

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

where,

p = proportion of interest

n = sample size

α = Level of significance (desired confidence)

$Z_{1-\frac{\alpha}{2}}$ = "Z value" for desired level of confidence

Normal approximation method works well when n is large, and p is neither very small nor very large. But for very small values of p it doesn't provide accurate results. Due to the inaccuracy of the normal approximation method, many statisticians started using the exact Clopper-Pearson method.

PhUSE 2013

CLOPPER PEARSON METHOD

Clopper-Pearson estimation method is based on the exact binomial distribution, and not a large sample normal approximation. When compared to Normal approximation method, this method is accurate when $np > 5$ or $n(1-p) > 5$ also the computation is possible when $p = 0$ or $p = 1$.

The formula for the confidence interval is given below:

$$\left(1 + \frac{n - x + 1}{x * F(1 - \frac{\alpha}{2}; 2x, 2(n - x + 1))} \right)^{-1} < p < \left(1 + \frac{n - x}{(x + 1) * F(\frac{\alpha}{2}; 2(x + 1), 2(n - x))} \right)^{-1}$$

Where x is the number of successes, n is the number of trials, p is the proportion of successes (x/n), α = the level of significance, $F(\alpha, b, c)$ = the specified α^{th} percentile of the F distribution with b and c degrees of freedom.

Because of a relationship between the cumulative binomial distribution and the beta distribution, the Clopper-Pearson interval has an alternate format that uses quantiles from the beta distribution.

$$B\left(\frac{\alpha}{2}; x, (n - x + 1)\right) < p < B\left(1 - \frac{\alpha}{2}; (x + 1), (n - x)\right)$$

Where x is the number of successes, n is the number of trials, and $B(p; v, w)$ is the p^{th} quantile from a beta distribution with shape parameters v and w . The lower bound is set to 0 when $x = 0$, and the upper bound is set to 1 when $x = n$.

There are other methods like Wilson Score method is available for the CI Computations, Clopper-Pearson interval is the most widely used because of its availability in almost all statistical software packages and considerably lower computational complexity when compared to the other methods.

This paper focuses on the use of binomial proportion and its confidence intervals in evaluating the overall incidence of adverse events in a clinical trial. It provides a brief idea on how to compute the Clopper-Pearson (exact) binomial confidence intervals on a sample data, using the PROC FREQ in SAS, and also provides some specific programming tips to avoid erroneous CI computations.

COMPUTATION OF BINOMIAL PROPORTION AND CI'S IN ADVERSE EVENT SUMMARIES

While assessing the overall incidence of adverse events in a clinical trial using the binomial proportion method, the 'Occurrence/Non Occurrence' of any adverse events (treatment emergent, serious, significant etc...) can be considered similar to the 'failure/success' outcomes in a binomial experiment. A typical summary table shell that is used in such trials, which contains the binomial proportions and their confidence intervals is as follows

Overall Incidence

	Treatment A (N=xx)		Treatment B (N=xx)	
	No of Subjects	Proportion of Subjects (95% CI)*	No of Subjects	Proportion of Subjects (95% CI)*
Any TEAE	xx	x.xxx (x.xxx-x.xxx)	xx	x.xxx (x.xxx-x.xxx)
Any Serious AE	xx	x.xxx (x.xxx-x.xxx)	xx	x.xxx (x.xxx-x.xxx)
Any Significant AE	xx	x.xxx (x.xxx-x.xxx)	xx	x.xxx (x.xxx-x.xxx)

*CIs obtained using Clopper –Pearson method.

Consider the below sample data which contain one observation per subject, with separate variables derived for each of the categories with values '0' (No) and '1' (Yes).

PhUSE 2013

'ae1' dataset

Subject	Treatment	ANYTEAE ¹	ANYSERAE ²	ANYSIGAE ³
001	A	1	1	1
002	A	0	0	0
003	A	1	1	0
004	A	0	0	0
005	A	0	0	0
006	A	1	1	1
007	A	0	0	0
008	A	1	1	1
009	A	1	1	0
010	A	1	1	1
011	B	0	0	0
012	B	1	0	0
013	B	1	1	0
014	B	1	1	0
015	B	0	0	0
016	B	1	1	0

¹ Any TEAE, ² Any Serious AE, ³ Any Significant AE

Suppose there are a total of 16 subjects in the trial, 10 subjects in TRT A and 6 subjects in TRT B. PROC FREQ in SAS can be used to compute the number of subjects, their proportions and the exact confidence intervals.

Following is a simple PROC FREQ code which provides the number of subjects, percentages and the exact CIs for the binomial proportions. Even if multiple variables can be used in the TABLE statement of the PROC FREQ, only one variable (ANYSIGAE) is used here for illustration purposes.

```
proc freq data=ae1;
  by treatment;
  tables anysigae/ nocum norow binomial;
  exact binomial;
run;
```

SAS output created from the PROC FREQ is as follows

For 'Treatment A', there are 4 subjects with 'Any Significant AE' as "Yes". Actual proportion of subjects is 0.40 and the CIs should be computed for that proportion. SAS output from the PROC FREQ indicates that the 95% CI's (0.2624, 0.8784) obtained are for the proportion of subjects with ANYSIGAE =0 (p=0.6). Hence the result is incorrect.

----- Treatment=A -----

The FREQ Procedure

ANYSIGAE	Frequency	Percent
0	6	60.00
1	4	40.00

Binomial Proportion for ANYSIGAE = 0

Proportion (P)	0.6000
ASE	0.1549
95% Lower Conf Limit	0.2964
95% Upper Conf Limit	0.9036
Exact Conf Limits	
95% Lower Conf Limit	0.2624
95% Upper Conf Limit	0.8784

PhUSE 2013

And for 'Treatment B', since there are no subjects with 'Any Significant AE' as "Yes", the proportion is zero. There also the 95% CIs obtained are for the (0.5407, 1.0000) the proportion of subjects with ANYSIGAE =0 (p=1).

----- Treatment=B -----

The FREQ Procedure

ANYSIGAE	Frequency	Percent
0	6	100.00

Binomial Proportion for ANYSIGAE = 0

Proportion (P)	1.0000
ASE	0.0000
95% Lower Conf Limit	1.0000
95% Upper Conf Limit	1.0000
Exact Conf Limits	
95% Lower Conf Limit	0.5407
95% Upper Conf Limit	1.0000

HOW TO OBTAIN CORRECT CONFIDENCE INTERVALS USING PROC FREQ?

Above example indicates that the CIs obtained from PROC FREQ might not always be correct. Possible reasons for these erroneous computations and some methods to compute the correct CIs are described below.

Reason 1: By default, PROC FREQ computes the CI for the lowest level of the variable which we use.

In the above example, when both the levels are available (eg. Treatment A, 4 subjects are with '0 (No)' and 6 subjects are with '1(Yes)'). So while computing the CIs, Lowest level i.e. '0' is used and the CI which is computed for the proportion of subjects with response as '0(No)'.

Solution:

To overcome this, it is advised to reset the level of variables in such a way that the response of interest should get the lowest level. To make the response '1(Yes)' lower to the other, we can reset the '0 (No)' to '2(No)'. Following code resets the '0' to '2' and then computes the CIs using PROC FREQ

```
data ae2;
    set ae1;
    if anysigae=0 then anysigae=2;
run;

proc freq data=ae2 ;
    by treatment;
    tables anysigae/ nocum norow binomial;
    exact binomial;
run;
```

Resulting SAS Output for Treatment A is as follows

PhUSE 2013

----- Treatment=A -----

The FREQ Procedure

ANYSIGAE	Frequency	Percent
1	4	40.00
2	6	60.00

Binomial Proportion for ANYSIGAE = 1

Proportion (P)	0.4000
ASE	0.1549
95% Lower Conf Limit	0.0964
95% Upper Conf Limit	0.7036
Exact Conf Limits	
95% Lower Conf Limit	0.1216
95% Upper Conf Limit	0.7376

Above SAS output indicates that the CI values for 'Treatment A' are now computed correctly for the proportion of subjects with ANYSIGAE =1 (p=0.4).and the values are changed to (0.1216, 0.7376).

Reason 2: Absence of a required response level (or when the proportion is zero).

For Treatment B, there are no subjects with ANYSIGAE=1 (ie. the resulting proportion is zero but level '1(Yes)' is missing in the dataset). In the absence of a lower level '1(Yes)', PROC freq considers the level '2(No)' as the lowest level and computes the confidence intervals for the proportion of subjects with ANYSIGAE =2 (p=1).

SAS Output for Treatment B.

----- Treatment=B -----

The FREQ Procedure

ANYSIGAE	Frequency	Percent
2	6	100.00

Binomial Proportion for ANYSIGAE = 2

Proportion (P)	1.0000
ASE	0.0000
95% Lower Conf Limit	1.0000
95% Upper Conf Limit	1.0000
Exact Conf Limits	
95% Lower Conf Limit	0.5407
95% Upper Conf Limit	1.0000

So it can be observed that even after resetting the '0' to '2', 95% CIs (0.5407, 1.0000) obtained from PROC FREQ are not correct.

Solution:

When a required level is missing, we need to add records to the dataset and then to make use the 'WEIGHT' statement in PROC FREQ to consider only the relevant records for the CI computations.

To add records with lowest level of the target variable, we can create a dataset which has the lowest level against all the treatments. For the above example, the following dataset can be used to add records to the existing ones.

PhUSE 2013

TRTLEVEL dataset

	Treatment	ANYSIGAE
1	A	1
2	B	1

If this dataset is merged with the original one, using Treatment and ANYSIGAE as BY variables, it should add a new record to the existing one with ANYSIGAE=1. If all the treatments contain at least one subject with ANYSIGAE=1, no records will be added.

Also a new variable need to be added (here it is wgt) to the dataset in such a way that the newly added records will get a value of '0' and the already existing records will get a value of '1'. This variable can then be used in the WEIGHT statement in PROC FREQ to compute CIs correctly (for ANYSIGAE=1), by taking the proportion as zero. SAS code that can be used to add new records and to create weight variable is as follows

```
data ae3;
  merge ae2(in=a) trtlelevel(in=b);
  by treatment anysigae;
  if b and not a then wgt=0;
  else wgt=1;
run;
```

Then the WEIGHT statement with 'zeroes' option can be used to compute the correct proportions and CIs. PROC FREQ code with WEIGHT statement is as follows.

```
proc freq data=ae3 ;
  by treatment;
  tables anysigae/ nocum norow binomial;
  weight wgt/zeroes;
  exact binomial;
run;
```

SAS Output for Treatment B.

----- Treatment=B -----		
The FREQ Procedure		
ANYSIGAE	Frequency	Percent
1	0	0.00
2	6	100.00
Binomial Proportion for ANYSIGAE = 1		
Proportion (P)	0.0000	
ASE	0.0000	
95% Lower Conf Limit	0.0000	
95% Upper Conf Limit	0.0000	
Exact Conf Limits		
95% Lower Conf Limit	0.0000	
95% Upper Conf Limit	0.4593	

It can be observed that the Binomial Proportion is now computed for ANYSIGAE=1 and the computed proportion is 0. Also this provides the correct t 95% CIs (0.0000, 0.4593) for the zero proportion.

It is clear from the above example that to obtain the correct CIs, we need to do the same steps for each of the categories (ie. separate datasets and PROC FREQ steps for ANYTEAE, ANYSERAE and ANYSIGAE) for which the proportions and CIs are to be computed.

When the number of categories for which the proportions and CIs needs to be computed are high, addition of new records and the creation of weight variables in each of the datasets may require some macro programming and that sometimes is difficult for the beginners to do.

PhUSE 2013

IS IT POSSIBLE TO COMPUTE BINOMIAL CI'S WITHOUT PROC FREQ?

When we have such a procedure like PROC FREQ in SAS which provides the binomial CIs directly, most of us tend to use the same without bothering about the alternate solutions available. But the above example explained the fact that extreme care should be taken while using the PROC FREQ to compute the exact binomial CIs.

In reality, if the number of subjects (x) in each of those categories and the total number of subjects in each treatment (N) are available, exact CIs can be computed easily with the use of a simple SAS data step. Also this method will be comparatively easy to use when the number of categories for which the proportions and CIs needs to be computed are high.

For the sample AE dataset used in the above examples, the number of subjects and total number of subjects are as follows

Treatment	Category	Number of Subjects (x)	Total Number of Subjects (N)
A	Any TEAE	6	10
A	Any Serious AE	3	10
A	Any Significant AE	4	10
B	Any TEAE	4	6
B	Any Serious AE	3	6
B	Any Significant AE	0	6

SAS Code to manually compute the binomial proportion and its 95% confidence interval when the 'n' and 'x' are given is as follows.

'aenum' dataset

	treatment	category	x	N
1	A	ANYTEAE	6	10
2	A	ANYSERAE	3	10
3	A	ANYSIGAE	4	10
4	B	ANYTEAE	4	6
5	B	ANYSERAE	3	6
6	B	ANYSIGAE	0	6

SAS Code to obtain the proportion and 95% CIs is as follows

```
data clop_ci;
  set aenum;
  p=round ((x/n),.0001);
  if p=0 then CI_LOW=0;
  if p=1 then CI_HIGH=1;
  if p ne 0 then CI_LOW=round((1-betainv(.975,(n-x+1),x)),.0001);
  if p ne 1 then CI_HIGH=round((1-betainv(.025,(n-x),x+1)),.0001);
run;
```

And the resulting output dataset is as follows

	treatment	category	x	N	p	CI_LOW	CI_HIGH
1	A	ANYTEAE	6	10	0.6	0.2624	0.8784
2	A	ANYSERAE	3	10	0.3	0.0667	0.6525
3	A	ANYSIGAE	4	10	0.4	0.1216	0.7376
4	B	ANYTEAE	4	6	0.6667	0.2228	0.9567
5	B	ANYSERAE	3	6	0.5	0.1181	0.8819
6	B	ANYSIGAE	0	6	0	0	0.4593

PhUSE 2013

In the above sample code, formula that uses quantiles from the beta distribution was used to compute the Clopper-Pearson interval. So the values 0.975 ($1-\alpha/2$) and .025 ($\alpha/2$) may vary with the chosen significance level. It can be observed that the CI_LOW and CI_HIGH values for 'ANYSIGAE' category are same to that of the correct CIs computed using PROC FREQ.

CONCLUSION

Computation of binomial proportion and exact confidence intervals is easy with PROC FREQ. But careful evaluation of the results obtained from a PROC FREQ showed that the use of an improperly structured data may result in erroneous computation of the CIs. This paper explains the use of appropriate options in PROC FREQ and provides some programming tips to restructure the data and to make it ready for the CI computations using PROC FREQ. Finally, by implementing the CI computation formula in a SAS data step, it suggests an alternate method to compute the exact CIs without the help of PROC FREQ.

REFERENCES

1. DasGupta, Anirban; Brown D, Lawrence; Cai, Tony; "Interval Estimation for a Binomial Proportion" ; Statistical Science, Vol. 16, No.2, pp. 101-117,2001.
2. Vollset, Stein; "Confidence Intervals for a Binomial Proportion"; Statistics in Medicine, Vol 12, pp. 809-824, 2006.
3. Dunnigan, Keith; "Confidence Interval Calculation for Binomial Proportions"; MWSUG, P08 – 2008.
4. Alan, Agresti; Coull, Brent; "Approximate is Better than 'Exact' for Interval Estimation of Binomial Proportions" The American Statistician, Vol. 52, No. 2, pp 119-126, 1998;

ACKNOWLEDGMENTS

I sincerely thank Mrs. Prajitha Nair, Project Manager - SAS Programming, Kreara Solutions Pvt. Ltd., for reviewing the paper and providing comments and suggestions.

CONTACT INFORMATION

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

Jose Abraham
Kreara Solutions Pvt. Ltd.
T4, 7th Floor, Thejaswini Building, Technopark
Thiruvananthapuram - 695 581, India.
Work Phone: +91-471-2527640
Email: jose@kreara.com
Web: www.kreara.com

Brand and product names are trademarks of their respective companies.