Clopper-Pearson Confidence Interval in SAS Using Proc Freq

Anik Chatterjee

Novartis, Hyderabad, India

NOVARTIS

Agenda :

- Introduction
  - Motivation
  - Case Study
- Idea of Confidence Interval
- Statistical Framework
- How SAS performs
  - Different options and resulting output
- Proposed Solution / Workaround
  - Individual Treatment arm
  - Treatment Difference

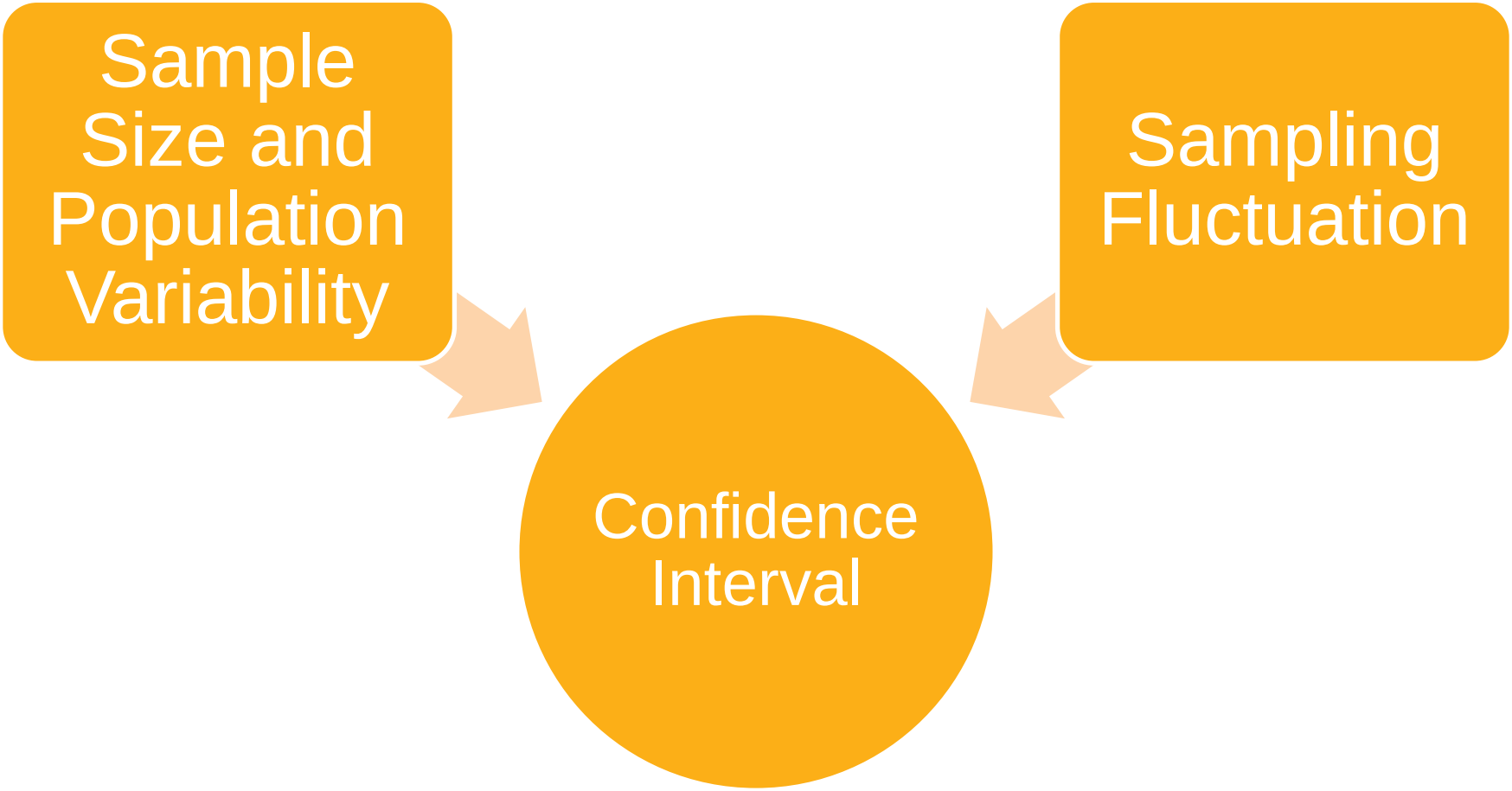
Introduction

Motivation and Case Study

- Subgroup Analysis in Clinical Trial
  - Objective of the subgroup analysis.
  - Parameters to be estimated
  - Challenges because of small sample size for a particular subgroup
- Case Study:
  - **Objective**: To study if there is a difference in treatment response by number of prior lines of treatment between the two treatment arms
  - **Parameters of interest**: Proportion of patients with Overall response(CR/PR) by prior lines of therapies as subgroups and treatment arm. In addition, to study the Difference in the proportions between the two treatment arms.

Idea of Confidence Interval

Idea of Confidence Interval



Idea of Confidence Interval

Sample and Population

- Population : Collection of **all possible respondents**.

We wish to know the proportion of **white respondent**.

Strategy: Will draw a random sample and measure the prop to estimate about population prop.

Sampling Fluctuations

Different sample from same population will give different result.

There will always be Sampling Fluctuation.

Inference

Drawing conclusion **about population based on the Sample**.

Idea of Confidence Interval

Different Sampling methods

- **SRS – WR / WOR**

  Simple random sampling – with replacement or without replacement.

- **Stratified Random Sampling**

  If we are aware, population is consists of different groups, then dividing population in groups and drawing random sample from each group to make al the groups equally representable in the sample.

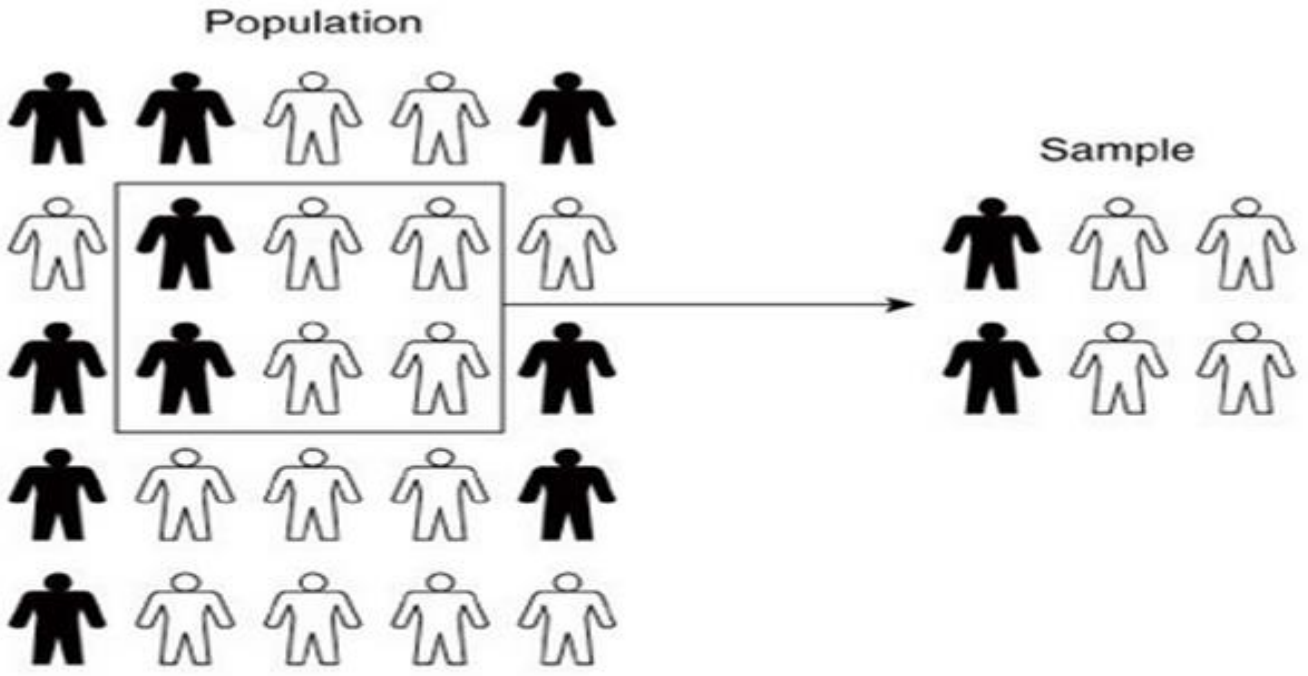
- **Cluster Sampling**

  If the population is homogeneous and widely spread, collecting sample might be time consuming. We may then draw a random sample to select which area to cover under consideration and then draw random sample form those selected area.

Idea of Confidence Interval

Sampling Fluctuations and Inference

The Proportion of White Respondents in a Population and in a Sample



Parameter

Proportion of white respondents in the population

$$\pi = \frac{15}{25} = .60$$

Statistic

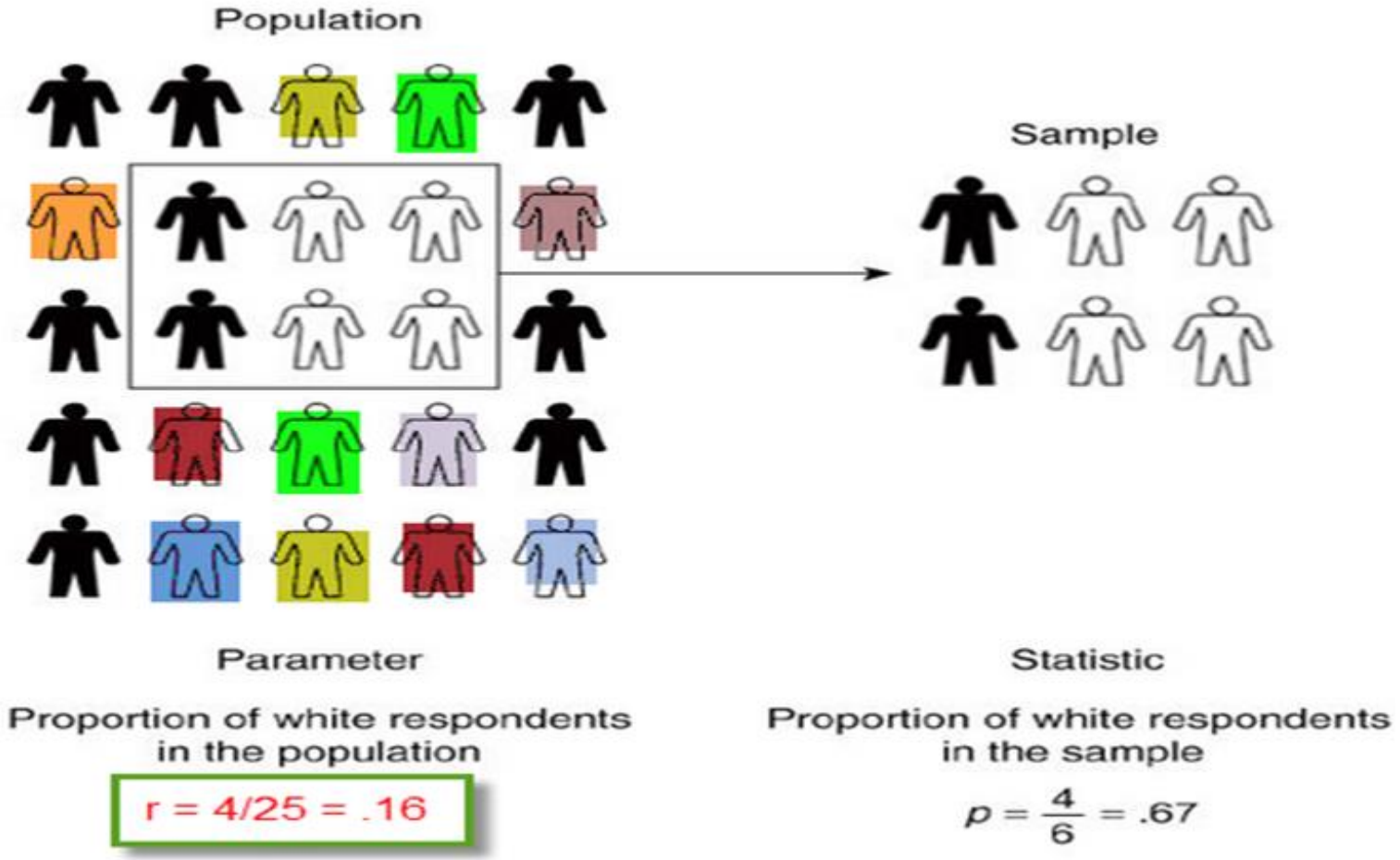
Proportion of white respondents in the sample

$$p = \frac{4}{6} = .67$$

Idea of Confidence Interval

Variation within the population

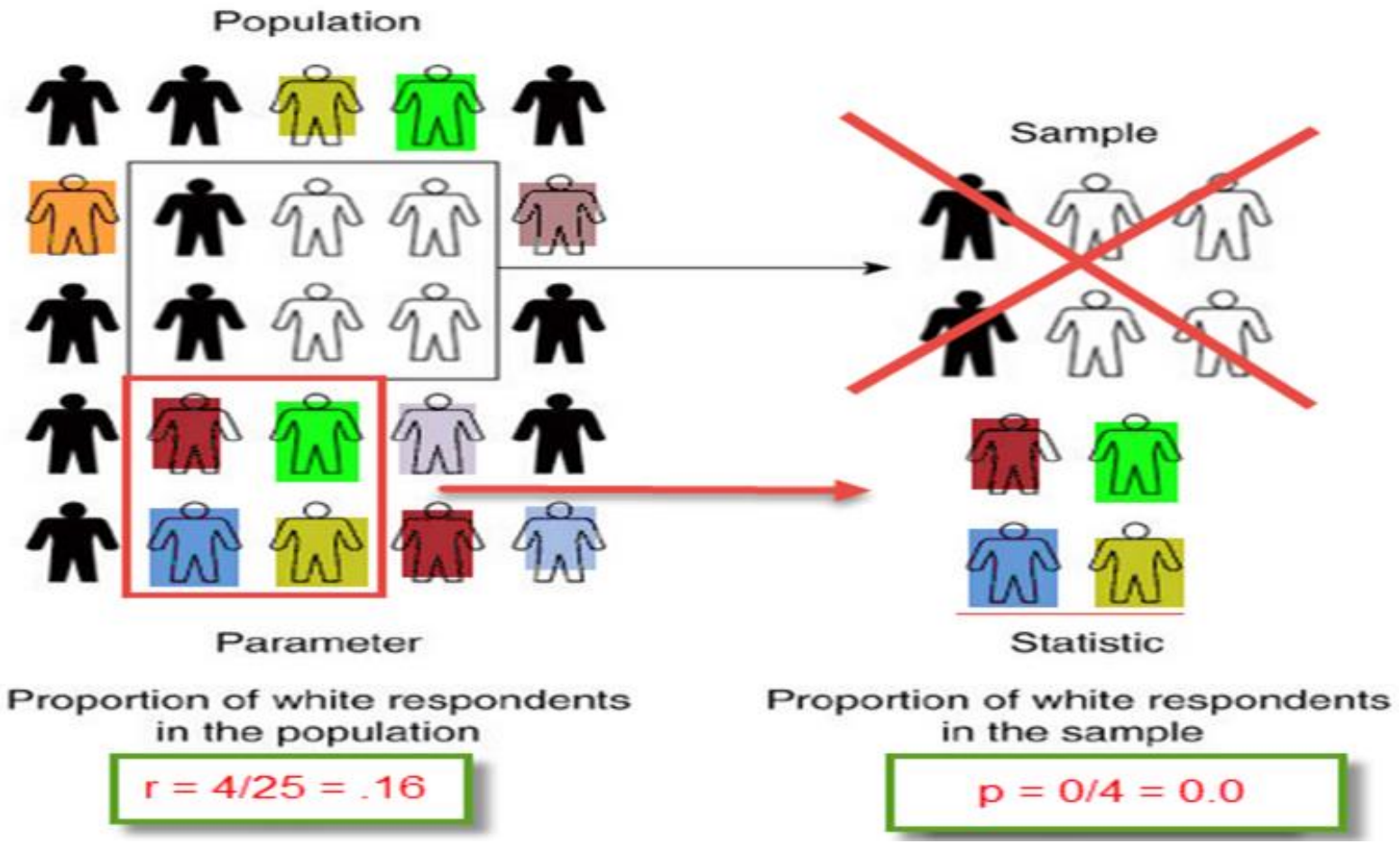
Less/More variation(P) => Less/More variation(S)



NOVARTIS

Idea of Confidence Interval

Variation due to Sample Size



NOVARTIS

Idea of Confidence Interval

What it tells?

- When we try to find an estimate of population parameter from sample statistic..... Good to provide an interval and to attach a confidence statement on the same.
- It tells.... How precise is the estimate.
- *What affects the width of CI ???*

Variation within the population of interest

Less # different respondents ➔ Each Sample will have less variation with other sample.

More # different respondents ➔ Each Sample will have more variation with other sample.

Size of the sample.

Idea of Confidence Interval

Effect on CI

Population	Sample	Estimate	CI
Less variation	All sample of similar kind	Good estimate close to population parameter	Smaller interval
More variation	More varied sample	Less confidence on result to estimate population parameter	Larger interval
	Small Sample Size	Less information	Bigger interval
	Big Sample Size	More information	Smaller interval

Idea of Confidence Interval

Which drug you prefer to take????

DRUG	Total Pat	SAE
A	10	2
B	100	20

n1	n	prop	lw	up
2	10	0.2	2.5210726327	55.609546231
20	100	0.2	12.66555521	29.184268909

Statistical Framework

Let us consider a random variable, $X_{ij}(i = \text{A and B}, j = 1,2)$

The number of occurrence of the adverse event corresponding to the treatment i for the j th subgroup.

Then X_{ij}'s are independently binomially distributed random variables with parameters N_{ij} and p_{ij}.

N_{ij} is the number of patients in the study with treatment i and subgroup j.

Here, our objective is to estimate p_{Aj}, p_{Bj} and $p_{Aj} - p_{Bj}$ with their Clopper Pearson Confidence interval.

Statistical Framework

Investigational drug is for patients who have progressed after at least one prior line of therapy at R-M Setting

Overall Response(CR/PR) by # of prior lines of therapy	TRTA (n/N) (%) 95% CI	TRTB (n/N) (%) 95% CI	Treatment Difference (%) 95% CI
Prior Lines of therapy in R-M Setting=1	32/46 69.6% (54.3 , 82.3)	15/50 30% (17.9 , 44.6)	39.6% (21.2, 58.0)
Prior Lines of therapy in R-M Setting=2	9/20 45% (23.1 , 68.5)	4/22 18.2% (5.2 , 40.3)	26.8% (-0.3, 53.9)
No prior lines of therapy at R-M Setting	0/5 0.0%	0/1 0.0%	0.0%
Others	3/9 33.3% (7.5 , 70.1)	0/8 0.0% (0.0, 36.9)	33.3% (2.5, 64.1)

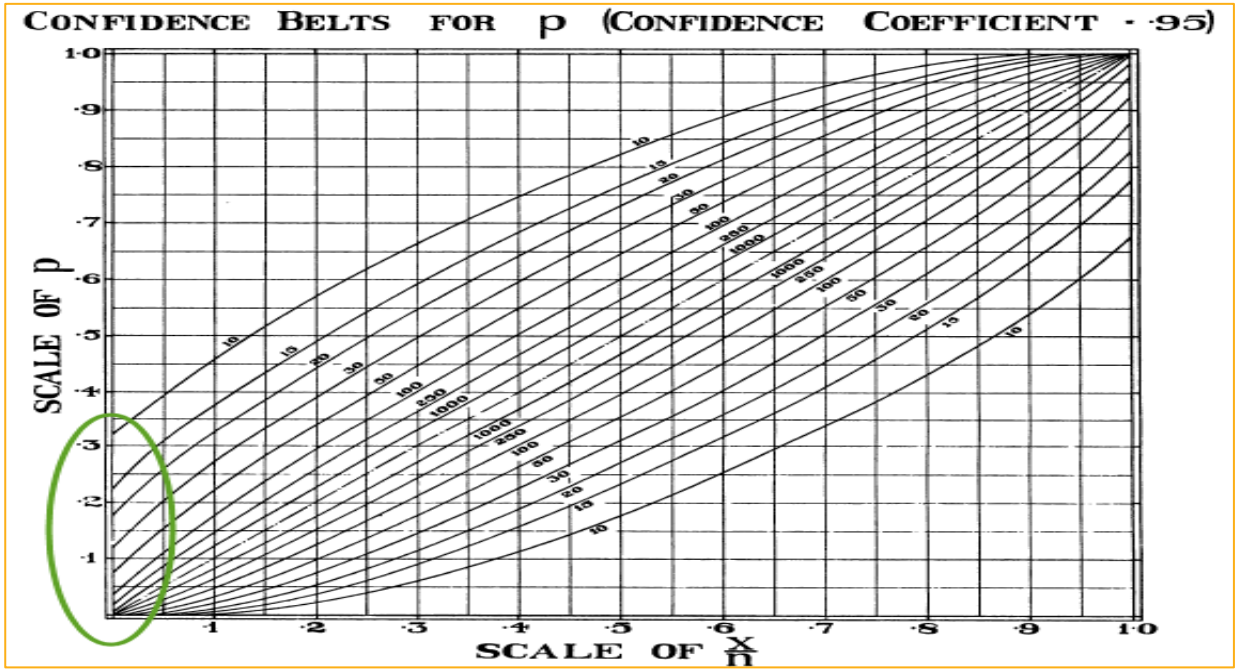
From the above table we observe $\hat{p}_{A2}$ and $\hat{p}_{B2}$ equals to zero.

Where $\hat{p}_{ij}$ is an estimate of p_{ij}.

Statistical Framework

The Clopper- Pearson method is capable of finding the confidence interval of $\hat{p}_{ij}$ when its estimate is zero.

Please see the screenshot below Clopper Pearson(1934).
The green circle shows the CI for p when estimate of p = 0 for n=10.



NOVARTIS

How SAS Performs

Different options and resulting output

- The Calculation of individual treatment arm and treatment difference should be available using the following options :

  PROC FREQ with **“exact” and “riskdiff”**

- When both the treatment arms are zero....

  Proc FREQ is unable to calculate CI for individual treatment arm and treatment difference as well.

Example .....

How SAS Performs

Different options and resulting output

- TRT : Treatment information (1 or 2).
- Event : 1 = occurrence of event and
  0 = non-occurrence of the event.
- Count : Frequency or no of patients for the occurrence and non-occurrence of the event for any particular treatment.

The dataset ASD1 the event has not occurred for none of the treatment arms, whereas for dataset ASD we have non-zero frequencies for all the categories.

ASD

trt	count	event
1	6	0
2	9	0
1	3	1
2	2	1

ASD1

	trt	count	event
1	1	6	0
2	2	9	0
3	1	0	1
4	2	0	1

How SAS Performs

Different options and resulting output

Screenshot of the SAS log that illustrates the issue of Proc Freq. The warning message is getting populated out of same SAS code for different scenario.

```
NOTE: Remote submit to SERVER commencing.
11425   ods output  riskdiffcol2=out2; *** output. will be the relevant dataset***;
11426   proc freq data=asd;
11427      tables trt*event/chisq exact riskdiff;     ***event: the variable with value 0/1 ***;
11428      weight count;                              *** count: number of patient with event=0/1 ***;
11429   run;

NOTE: The data set WORK.OUT2 has 4 observations and 9 variables.
NOTE: There were 4 observations read from the data set WORK.ASD.
NOTE: The PROCEDURE FREQ printed pages 8-10.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.03 seconds
      cpu time            0.00 seconds


11430   *ods trace off;
11431   ods listing;
11432
11433   ods output  riskdiffcol2=out3; *** output. will be the relevant dataset***;
11434   proc freq data=asd1;
11435      tables trt*event/chisq exact riskdiff;     ***event: the variable with value 0/1 ***;
11436      weight count;                              *** count: number of patient with event=0/1 ***;
11437   run;

NOTE: No statistics are computed for trt * event since event has less than 2 nonmissing levels.
WARNING: Output 'riskdiffcol2' was not created.  Make sure that the output object name, label, or path is spelled
         correctly.  Also, verify that the appropriate procedure options are used to produce the requested output
         object.  For example, verify that the NOPRINT option is not used.
NOTE: There were 4 observations read from the data set WORK.ASD1.
NOTE: The PROCEDURE FREQ printed page 11.
NOTE: PROCEDURE FREQ used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


11438   *ods trace off;
11439   ods listing;
NOTE: Remote submit to SERVER complete.
```

How SAS Performs

Different options and resulting output

The following screenshot from SAS webpage describes how the CI for treatment difference is calculated and in principle this should be available as well for the above mentioned cases.

PROC FREQ computes the confidence limits as follows. The risk difference is defined as the difference between the row 1 and row 2 risks (proportions), $d = p_1 - p_2$, and n_1 and n_2 denote the row totals of the 2×2 table. The joint probability function for the table can be expressed in terms of the table cell frequencies, the risk difference, and the nuisance parameter p_2 as

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

The $100(1-\alpha/2)\%$ confidence limits for the risk difference are computed as

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

where

$$P_U(d_*) = \sup_{p_2}\Big(\sum_{A.T(a)\geq t_0} f(n_{11},n_{21};n_1,n_2,d_*,p_2)\Big)$$

$$P_L(d_*) = \sup_{p_2}\Big(\sum_{A.T(a)\leq t_0} f(n_{11},n_{21};n_1,n_2,d_*,p_2)\Big)$$

The set A includes all 2×2 tables with row sums equal to n_1 and n_2, and $T(a)$ denotes the value of the test statistic for table a in A. To compute $P_U(d_*)$, the sum includes probabilities of those tables for which $(T(a) \geq t_0)$, where t_0 is the value of the test statistic for the observed table. For a fixed value of d_*, $P_U(d_*)$ is taken to be the maximum sum over all possible values of p_2.

Proposed Solution

Individual Treatment arm

Using Proc FREQ

trt	x	ct
1	0	4
2	0	6
1	1	0
2	1	0

trt	Table	Name1	Label1	cValue1
1	Table x	_BIN_	Proportion	0.0000
1	Table x			
1	Table x		Exact Conf Limits	
1	Table x	XL_BIN	95% Lower Conf Limit	0.0000
1	Table x	XU_BIN	95% Upper Conf Limit	0.6024
2	Table x	_BIN_	Proportion	0.0000
2	Table x			
2	Table x		Exact Conf Limits	
2	Table x	XL_BIN	95% Lower Conf Limit	0.0000
2	Table x	XU_BIN	95% Upper Conf Limit	0.4593

```
data test;
  trt=1; x=1; ct=0; output;
  trt=1; x=0; ct=4; output;
  trt=2; x=1; ct=0; output;
  trt=2; x=0; ct=6; output;
run;

proc freq data=test;
  weight ct/zeros;
  tables x/binomial(level='1');
  by trt;
run;
```

Proposed Solution

Individual Treatment arm

Using Data Step / SQL

Using the **FINV** function that returns the p[th] quantile from the F distribution.

FINV(p , ndf , ddf , nc)

p : numeric probability (0 <= p <= 1)
ndf : numerator degrees of freedom
ddf : denominator degrees of freedom
nc : non-centrality parameter(optional)

```
data asd;
   n1 = 0;n = 4;output;
   n1 = 0;n = 6;output;
run;

data asdf;
set asd;
/*Clopper-Pearson CI for individual treatment*/
   if n1=0 then lw = 0;
     else lw = ((1 + ((n-n1+1)/(finv(0.05/2,2*n1,2*(n-n1+1))*n1)) )**(-1));

   if n1=n then up = 1;
     else up = (1 + ((n-n1)/(finv(1-(0.05/2),2*(n1+1),2*(n-n1))*(n1+1))))**(-1);
run;
```

n1	n	lw	up
0	4	0	0.6024
0	6	0	0.4593

Proposed Solution

Treatment Difference

Newcombe Interval

- A good confidence interval for the difference of proportions will produce an appropriate distribution of coverage probabilities.
- The standard Wald interval that SAS produces with PROC FREQ and RISKDIFF produces a zero-width interval when both proportions in a data set are zero.
- The Newcombe interval, based on the two individual Wilson intervals, has excellent coverage probabilities and outperforms nearly all other methods.
- It is computationally simple and will never give any aberrations (like negative proportions or zero-width intervals).

Proposed Solution

Treatment Difference

```
data test;
  trt=1; x=1; ct=0; output;
  trt=1; x=0; ct=4; output;
  trt=2; x=1; ct=0; output;
  trt=2; x=0; ct=6; output;
run;

proc freq data=test;
  weight ct/zeros;
  tables x/binomial(wilson level='1');
  by trt;
  output out=wilson binomial;
run;

data w1 (keep=_bin_ l_w_bin u_w_bin rename=(_bin_=p1 l_w_bin=l1 u_w_bin=u1))
     w2 (keep=_bin_ l_w_bin u_w_bin rename=(_bin_=p2 l_w_bin=l2 u_w_bin=u2));
  set wilson;
  if trt=1 then output w1;
  if trt=2 then output w2;
run;
```

trt	x	ct
1	1	0
1	0	4
2	1	0
2	0	6

Proposed Solution

Treatment Difference

```
data wilson2;
  merge w1 w2;
  est = p1 - p2;
  low = est - sqrt((p1-l1)**2 + (u2-p2)**2);
  upp = est + sqrt((u1-p1)**2 + (p2-l2)**2);
  call symput('est', put(est, 6.4));
  call symput('low', put(low, 6.4));
  call symput('upp', put(upp, 6.4));
run;
%put Difference is &est with 95% CI [&low., &upp];
```

p1	l1	u1	p2	l2	u2	est	low	upp
0	0	0.4898908365	0	0	0.3903342879	0	-0.390334288	0.4898908365

References :

- *The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial*

  C. J. Clopper; E. S. Pearson

- *INTERVAL ESTIMATION FOR THE DIFFERENCE BETWEEN INDEPENDENT PROPORTIONS: COMPARISON OF ELEVEN METHODS*

  ROBERT G. NEWCOMBE

Suggestion and Discussion

Anik Chatterjee
Principal Programmer
Novartis, India

Contact
anik.chatterjee@novartis.com
anikchatterjee@gmail.com

NOVARTIS