

### ***Alternative Method to Bootstrap to Obtain Confidence Interval for Relative Treatment Difference***

Ruji Yao, Amarjot Kaur, Qing Li, Anjela Tzontcheva

*Merck & Co., Inc., NJ, USA;*

In a clinical trial, many times a treatment effect is expressed as ratio of two random variables, such as treatment difference relative to placebo. This presents a problem of computing the confidence interval (CI) where there is no closed form expression for a ratio of two random variables. Bootstrap is a well accepted method to obtain such confidence interval since the method is non-parametric and it is well defined for implementation. But it takes time for bootstrap. In this presentation, we present an alternative method that notably shortens the time of computation under few reasonable assumptions. This method is based on the concept of Fieller's theorem and assume that two estimated treatment effects follow a bivariate normal distribution and use numerical method directly obtain the frequency distribution using histogram of the ratio and then calculate the confidence interval from the histogram. The proposed method extends Fieller's method and can be utilized to calculate a confidence interval of any function of two random variables in addition to ratio, e.g., relative treatment effect in terms of two medians with normal assumption on the ranks of sample median. This results from the real trial data are the very similar to as bootstrap method, but confidence interval can always be obtained with relatively much shorter duration.

#### **Introduction**

**Bootstrap example** of CI of relative treatment effect to placebo, which is defined as  $100 \times (\text{effect of active} - \text{effect of placebo}) / (\text{effect of placebo})$ , of 3 primary and key secondary variables of ANOVA models from a real trial. Each has 10,000 bootstrap runs for CI of relative treatment difference.

To get a **unique result**, both random seeds and number of bootstrap runs have to be prespecified in the protocol.

<b>variable</b>	<b>Relative effect</b>	<b>Upper Bound</b>	<b>Upper Bound</b>	<b>Time of bootstrap</b>
TCS	-38.3	-46.0	-29.7	4658 sec
DSS	-35.4	-43.2	-26.1	4634 sec
DMS	-42.3	-52.2	-31.2	4644 sec

**Idea 1. If effects are in LS means**

If effects are in LS means, then the inference of treatment difference of 2 estimated means, are usually based on the bivariate normal distribution with parameters from model fitting. This is well known T-distribution inference.

For other functions of these 2 estimated means, like relative difference, it may not easily get an analytic formula. Our idea is to construct a histogram of interesting function under the same assumption of bivariate normal distribution.

Actually, Fieller’s formula can be used for CI of relative difference with the same assumption of bivariate normal distribution, but our histogram method is more straight forward and can be extended to other functions of 2 random variables.

**Method 1 Histogram Method for Relative Difference**

Assume the bivariate normal distribution from model fitting has mean vector  $(LSMa, LSMp)$  and standard error  $stderr_a, stderr_p$ , correlation coefficient  $\rho$ .

start from 100 x 100 grid on (0,1) x(0,1) uniform distribution.

(one can also start from 1000 x 1000 grid.)

Step 1. for each a from 0.005, 0.015, ..., 0.995, get

$$ma = probit(a) * stderr_a + LSMa$$

Step 2. for each p from 0.005, 0.015, ..., 0.995, get

$$mp = probit(p) * stderr_p * (1 - \rho^2) + LSMp + \rho * \left( \frac{stderr_p}{stderr_a} \right) * (ma - LSMa)$$

Step 3. get histogram of relative treatment effect as

$$100 * \frac{ma - mp}{mp} = 100 * \left( \frac{ma}{mp} - 1 \right)$$

**Note: One can create a histogram of any function of  $(ma, mp)$**

Step 4. Sort these 10,000 of relative treatment effect values and use 250<sup>th</sup> and 9750<sup>th</sup> values for 95%CI

**Result: 95% CI by difference method**

<b>variable</b>	<b>Relative effect</b>	<b>Relative effect</b>	<b>Lower Bound</b>	<b>Upper Bound</b>	<b>Time (sec)</b>
TCS	-38.3	bootstrap	-46.0	-29.7	4658
		Fieller formula	-46.7	-29.3	0
		histogram	-46.5	-29.5	0.36
DSS	-35.4	bootstrap	-43.2	-26.1	4634
		Fieller formula	-44.0	-26.1	0
		histogram	-43.8	-26.3	0.38
DMS	-42.3	bootstrap	-52.2	-31.2	4644
		Fieller formula	-53.0	-30.6	0
		histogram	-52.8	-31.0	0.44

**Use Histogram method to obtain 95% CI of treatment difference in proportions with logistic regression (SAS PROC LOGISTIC)**

Logistic regression can be used for model binary data with covariates including treatment and other interesting variables. From model, we can have logit estimate of treatment effect of active and placebo, and their difference in logit scale. It is easy to use inverse logit function to get estimated treatment effect in proportion. But there is no easy way to get the 95% CI of treatment difference in proportion.

Since the estimated treatment effect of active and placebo in logit are both normal distributed and all estimated parameters are available to define this bivariate normal distribution, we can simply replace the formula in Step 3 of previous section from  $100 * \frac{ma-mp}{mp}$  to  $(\text{inverse logit}(ma) - \text{inverse logit}(mp))$  and without any changes in Step 1, 2 and 4.

**Example: 95% CI of treatment difference in proportion for logistic regression**

**95% CI for treatment difference in proportion ( 0.16062) :**

**Bootstrap (10,000 runs): ( 0.09634, 0.22614) time used: 8536 sec**

**Histogram method: ( 0.09649, 0.22305) time used: 4 sec**

Results of SAS PROC LOGISTIC			
treatment	estimate in logit	stderr	inverse logit in proportion
Active	0.337	0.0966	0.58356
Placebo	-0.311	0.0944	0.42294
Difference	0.648	0.1329	0.16062

**Idea 2. If effects are in medians**

Suppose a bootstrap of N runs was used to get 95% CI for treatment difference in medians (not median of differences in H-L method). For each treatment group, a histogram of these N median values can be used for inference of any function of estimated medians. Note: 1. the histogram is a function of ranks of the data set, not its values. 2. the histogram on ranks is symmetric about sample median, which is true in general. 3. Normal approximation of this bell-shaped histogram: we use the formula of CI for median in survival analysis, normal  $(\mu, \sigma^2) = (0.5, \frac{0.5^2}{n})$ .

(ref: SAS manual of PROC LIFETEST)

**Method 2 Histogram Method for Median Difference**

Start from active group: on a (0,1) uniform distribution.

1. for each  $a$  from 0.005, 0.015, ..., 0.995, get a mapping into normal  $(\mu, \sigma^2) = (0.5, \frac{0.5^2}{n})$  as

$$p = \text{probit}(a) * 0.5/\sqrt{n} + 0.5$$

2. get their rank by multiple sample size  $n$

$$\begin{aligned} \text{rank}(p) &= n * (\text{probit}(a) * \frac{0.5}{\sqrt{n}} + 0.5) \\ &= \text{probit}(a) * \frac{\sqrt{n}}{2} + \frac{n}{2} \end{aligned}$$

3. Get 100 median values from the original data set by  $\text{rank}(p)$ ; Call it  $\{a\}$ , say  $a_{00}, a_{01}, \dots, a_{99}$ .

4. Do the same for placebo;

Call it  $\{p\}$ , say  $p_{00}, p_{01}, \dots, p_{99}$ .

These are the histograms of re-sample medians.

5. Get numerical histogram of difference of median  $\{a\} - \{p\}$ ,

$(a_{00}-p_{00}, a_{00}-p_{01}, \dots, a_{99}-p_{98}, a_{99}-p_{99})$

from these 100\*100 points and use it to calculate CI.

6. we can also get CI for relative treatment difference of median  $100(\{a\} - \{p\})/\{p\}$  in a similar way.

Below are examples of 95% CI of difference and relative difference in medians. TCS data is used..

<b>95% CI of difference to placebo in median</b>			
Method	Lower Bound	Upper Bound	Time (sec)
Bootstrap (10,000run)	-3.311	-1.638	4157
Histogram method	-3.267	-1.643	0.82
<b>95% CI of relative difference to placebo in median</b>			
Method	Lower Bound	Upper Bound	Time (sec)
Bootstrap (10,000run)	-59.6	-37.3	4157
Histogram method	-59.57	-37.27	0.82

**Summary:**

We proposed a new method based on the histogram method to calculate CI for ratios of two random variables, such as relative treatment effect in LS means or in medians based on the assumption of a bivariate normal distribution. The proposed method took very short computational time with results very similar to that based on bootstrap method for several real data sets. A good normal approximation is critical to success of this method, which is usually satisfied in a pivotal clinical trial that are adequately powered for statistical inference based on normal assumptions.