

The Combining Confidence Distribution Method to the Behrens-Fisher Problem

Wenyong Tao, Wanzhou Ye

School of Science, Shanghai University, Shanghai, China
Email: twyxjr2012@163.com

Received 14 June 2016; accepted 12 July 2016; published 15 July 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Based on the Confidence Distribution method to the Behrens-Fisher problem, we consider two approaches of combining Confidence Distributions: *P* Combination and *AN* Combination to solve the Behrens-Fisher problem. Firstly, we provide some Confidence Distributions to the Behrens-Fisher problem, and then we give the Confidence Distribution method to the Behrens-Fisher problem. Finally, we compare the “combination” and the “single” through the numerical simulation.

Keywords

Behrens-Fisher Problem, Combining Confidence Distribution, Interval Estimation Component

1. Introduction

1.1. Behrens-Fisher Problem

Let $X_{i1}, X_{i2}, \dots, X_{in_i}$, $i = 1, 2$, be i.i.d samples from two normal populations $N(\mu_i, \sigma_i^2)$, $i = 1, 2$. Four parameters $\mu_1, \mu_2, \sigma_1, \sigma_2$ are assumed to be unknown and not necessarily equal. Behrens-Fisher problem is to give the interval estimation of the parameter $\theta = \mu_1 - \mu_2$.

In the case of 1) $\sigma_1^2 = \sigma_2^2$ but unknown; 2) $\sigma_1^2 / \sigma_2^2 = c$ and $n_1 = n_2$, we can use frequentist approach to solve Behrens-Fisher problem; In general case, we usually use of large sample theory to find the approximate confidence interval [1].

1.2. Confidence Distribution

In Bayesian inference, researchers typically rely on a posterior distribution to make inference on a parameter of interest, where the posterior is often viewed as a “distribution estimator” [2] for the parameter.

Confidence Distribution is one such a “distribution estimator” that can be defined and interpreted in a fre-

quantist framework, in which the parameter is a fixed and non-random quantity. The concept of confidence distribution has a long history. The following definition is proposed and utilized in Schweder & Hjort (2002) [3] and Singh et al. (2005, 2007) [4] [5].

Definition 1.1: Given: Θ is the parameter space of the unknown parameter of interest θ ; X is the sample space corresponding to sample data $x = x_1, \dots, x_n$. We called the function $H_n(\cdot) = H_n(x, \cdot)$ a confidence distribution (CD) for a parameter θ , if

- 1) For each given $x \in X$, $H_n(\cdot)$ is a cumulative distribution function on Θ ;
- 2) At the true parameter value $\theta = \theta_0$, $H_n(\theta_0) \equiv H_n(x, \theta_0)$, as a function of the sample, follows the uniform distribution $U[0,1]$.

Also, the function $H_n(\cdot)$ is an asymptotic confidence distribution, if the $U[0,1]$ requirement is true only asymptotically.

And, we got a theorem in the author's another paper [6]:

Theorem 1.1: If for each given $x \in X$, $H_n(\cdot)$ is a cumulative distribution function on Θ , then we can get that at the true parameter value $\theta = \theta_0$, $H_n(\theta_0) \equiv H_n(x, \theta_0)$, as a function of the sample, follows the uniform distribution $U[0,1]$.

According to Theorem 1.1, we can see that if the $H_n(\cdot)$ meet the condition 1) in the Definition 1.1; then we can proof $H_n(\cdot)$ meet the condition 2) in the Definition 1.1. So $H_n(\cdot)$ is a confidence distribution.

To put it simply, Confidence Distribution is a distribution of the parameter, we can know almost all of the information of the parameter. But methods to the construction of the Confidence Distribution are not unique, so we can get different Confidence Distributions and then find the optimal one.

1.3. Some Confidence Distributions to the Behrens-Fisher problem

According to the conclusion of the author's another paper: in both small sample size and big sample size, the effectiveness of WS and CA are relatively close, but CA is a little better than WS in the optimality; And then we consider the forms of WS and CA are relatively simple.

So we choose WS and CA to combine Confidence Distribution. The following are the conclusions of WS and CA.

- 1) WS Distribution's density function:

$$f_{ws}(\theta) = \frac{1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{[\theta - (\bar{X}_1 - \bar{X}_2)]^2}{k}\right)^{-\frac{k+1}{2}}$$

where $s_i = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 / (n_i - 1)$.

- 2) CA Distribution's density function:

$$f_{ca}(\theta) = \frac{1}{\sqrt{2\pi} \left[(\hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2) \right]^{1/2}} \exp \left\{ -\frac{[\theta - (\bar{X}_1 - \bar{X}_2)]^2}{2 \left[\hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2 \right]} \right\}$$

where parameters in $\hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2$ are the solutions of the following equations:

$$\begin{cases} \sigma_1^2 = \left(\frac{s_1^2 (n_1 - 1)}{n_1} \right) + (n_2 \sigma_1^2)^2 (\bar{X}_1 - \bar{X}_2) / (n_2 \sigma_1^2 + n_1 \sigma_2^2)^2 \\ \sigma_2^2 = \left(\frac{s_2^2 (n_2 - 1)}{n_2} \right) + (n_1 \sigma_2^2)^2 (\bar{X}_1 - \bar{X}_2) / (n_1 \sigma_2^2 + n_2 \sigma_1^2)^2 \end{cases}$$

2. Combination

The notion of a Confidence Distribution is attractive for the purpose of combining information. The main reasons are that there is a wealth of information on θ inside a Confidence Distribution, the concept of Confidence

Distribution is quite broad, and the Confidence Distributions are relatively easy to construct and interpret.

2.1. P Combination

Multiplying likelihood functions from independent sources constitutes a standard method for combining parametric information. Naturally, this suggests multiplying CD densities and normalizing to possibly derive combined CDs as follows [3]:

$$H_p(\theta) = \int_{(-\infty, \theta) \cap \Theta} h^*(y) dy / \int_{(-\infty, +\infty) \cap \Theta} h^*(y) dy$$

where $h^*(y) = \prod_{i=1}^L h_i(y)$ and $h_i(y)$ are Confidence Distribution densities from L independent studies.

Now, we use the P Combination to combine WS and CA:

$$H_p(\theta) = \frac{\int_{(-\infty, \theta)} f_{ws}(y) \cdot f_{ca}(y) dy}{\int_{(-\infty, +\infty)} f_{ws}(y) \cdot f_{ca}(y) dy}$$

Theorem 2.1: $H_p(\theta)$ is a Confidence Distribution of $\theta = \mu_1 - \mu_2$.

Proof: Obviously, the $H_p(\theta)$ is a cumulative distribution function of θ , meet the condition 1) in definition 1.1; according to theorem 1.1, meet the condition 2) in definition 1.1. So $H_p(\theta)$ is a Confidence Distribution of $\theta = \mu_1 - \mu_2$.

2.2. AN Combination

Then we consider an asymptotic normality method based on asymptotic Confidence Distributions [7]:

$$H_{AN}(\theta) = 1 - \Phi\left(\left[\sum_{i=1}^L \frac{n_i}{V_i}\right]^{1/2} (\hat{\theta}_c - \theta)\right)$$

where $\hat{\theta}_c = \left(\sum_{i=1}^L \frac{n_i}{V_i} T_i\right) / \sum_{i=1}^L \left(\frac{n_i}{V_i}\right)$, here we let T_i are means of normal samples,

$$V_i = s_i = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 / (n_i - 1).$$

Now, we use the AN Combination to combine WS and CA:

$$H_{AN}(\theta) = 1 - \Phi\left(\left[\frac{n_{ws}}{s_{ws}^2} + \frac{n_{ca}}{s_{ca}^2}\right]^{1/2} \left[\left(\frac{n_{ws}}{s_{ws}^2} \cdot \mu_{ws} + \frac{n_{ca}}{s_{ca}^2} \cdot \mu_{ca}\right) / \left(\frac{n_{ws}}{s_{ws}^2} + \frac{n_{ca}}{s_{ca}^2}\right) - \theta\right]\right)$$

where $n_{ws} = n_{ca} = n_1 + n_2$.

Theorem 2.2: $H_{AN}(\theta)$ is a asymptotic Confidence Distribution of $\theta = \mu_1 - \mu_2$.

Proof: Let $\xi_i = \sqrt{n_i} \frac{T_i - \theta}{V_i}$, The normality based asymptotic Confidence Distribution is $H_i(\theta) = 1 - \Phi(\xi_i)$,

with asymptotic Confidence Distribution density $h_i(\theta) = \frac{\sqrt{n_i}}{V_i} \phi(\xi_i)$, Consider the combined function $H_p(\cdot)$

with input asymptotic Confidence Distribution densities $h_i(\cdot)$, so $H_{AN}(\theta)$ is a asymptotic Confidence Distribution of $\theta = \mu_1 - \mu_2$.

3. Numerical Simulation

3.1. Effectiveness

First of all, we need to consider the effectiveness of the Confidence Distribution in Behrens-Fisher problem. Here, we define the effectiveness of the Confidence Distribution:

$$\eta = P\left(H_{\alpha/2} < \bar{X}_1 - \bar{X}_2 < H_{1-\alpha/2}\right)$$

In this problem, we have a very small sample. In the numerical simulation, we define:

$$\hat{\eta} = \sum_{i=1}^n \frac{I\left(H_{\alpha/2} < \bar{X}_1^i - \bar{X}_2^i < H_{1-\alpha/2}\right)}{n}$$

where, I is a indicative function. The more $\hat{\eta}$ is close to α , the more Confidence Distribution is efficient.

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your journal for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

3.2. Optimality

Both $H_1(\cdot)$ and $H_2(\cdot)$ are Confidence Distribution of θ , if $H_{1,1-\alpha/2} \leq H_{2,1-\alpha/2}$ and $H_{1,\alpha/2} \geq H_{2,\alpha/2}$, then we call $H_1(\cdot)$ is better than $H_2(\cdot)$ at the confidence level on $1 - \alpha$ [5] [8].

3.3. Numerical Simulation

In the case of similar effectiveness, we consider the length of the confidence interval, the shorter length of the confidence interval corresponding Confidence Distribution is optimum.

According to the result of numerical simulation (Table 1), we can see:

Table 1. Under the condition of $\alpha = 0.05$, the effectiveness $\hat{\eta}$ of the different confidence distribution.

| (n_1, n_2) | Confidence Distribution | the effectiveness $\hat{\eta}$ under different σ_1^2/σ_2^2 | | | | | |
|--------------|-------------------------|--|---------|---------|---------|---------|---------|
| | | 0.1 | 0.5 | 1.0 | 2.0 | 5.0 | 10.0 |
| (10, 5) | WS | 0.05013 | 0.05176 | 0.06225 | 0.07228 | 0.07980 | 0.08321 |
| | CA | 0.04623 | 0.04372 | 0.04188 | 0.03812 | 0.03595 | 0.02665 |
| | PC | 0.04760 | 0.04613 | 0.05138 | 0.05333 | 0.05874 | 0.06078 |
| | AN | 0.04829 | 0.04712 | 0.04936 | 0.04781 | 0.04525 | 0.04357 |
| (10, 10) | WS | 0.05188 | 0.04839 | 0.04753 | 0.04914 | 0.05024 | 0.05241 |
| | CA | 0.04256 | 0.04820 | 0.04873 | 0.04724 | 0.04582 | 0.04413 |
| | PC | 0.04691 | 0.04835 | 0.04872 | 0.04860 | 0.04602 | 0.04608 |
| | ANC | 0.04722 | 0.04845 | 0.04860 | 0.04724 | 0.04682 | 0.04613 |
| (10, 25) | WS | 0.05027 | 0.04988 | 0.04931 | 0.04601 | 0.04079 | 0.03856 |
| | CA | 0.04182 | 0.04590 | 0.04836 | 0.04993 | 0.04918 | 0.04825 |
| | PC | 0.04658 | 0.04781 | 0.04944 | 0.04894 | 0.04825 | 0.04829 |
| | ANC | 0.04711 | 0.04861 | 0.04920 | 0.04820 | 0.04861 | 0.04853 |
| (25, 25) | WS | 0.04943 | 0.04893 | 0.04977 | 0.04971 | 0.05103 | 0.05054 |
| | CA | 0.04927 | 0.04982 | 0.04979 | 0.04930 | 0.04791 | 0.04632 |
| | PC | 0.04951 | 0.04915 | 0.04957 | 0.04970 | 0.04884 | 0.04807 |
| | ANC | 0.04911 | 0.04900 | 0.04949 | 0.04931 | 0.04916 | 0.04835 |
| (25, 50) | WS | 0.05004 | 0.05028 | 0.04868 | 0.04860 | 0.04694 | 0.04606 |
| | CA | 0.04837 | 0.04962 | 0.04938 | 0.04987 | 0.04880 | 0.04973 |
| | PC | 0.04922 | 0.04929 | 0.04938 | 0.04919 | 0.04872 | 0.04817 |
| | ANC | 0.04902 | 0.04983 | 0.04912 | 0.04909 | 0.04885 | 0.04859 |
| (50, 50) | WS | 0.04986 | 0.05027 | 0.04988 | 0.04900 | 0.04970 | 0.04990 |
| | CA | 0.04928 | 0.04904 | 0.04978 | 0.04885 | 0.04903 | 0.04820 |
| | PC | 0.04926 | 0.04921 | 0.04942 | 0.04966 | 0.04948 | 0.04822 |
| | ANC | 0.04987 | 0.04912 | 0.04960 | 0.04873 | 0.04915 | 0.04884 |

Where PC is for P Combination; ANC is for AN Combination.

- 1) With the increase of sample size, the effectiveness of each Confidence Distribution and combining Confidence Distribution increase.
- 2) The effectiveness of PC and ANC is better than WS and CA.

4. Conclusion

We use two different methods to combine Confidence Distributions. Through the numerical simulation we can find the optimal Confidence Distribution. In small sample size, the effectiveness of PC and ANC is better than WS and CA (PNC is a little better than PC); in the relatively big sample size, WS and CA are also effective. So, this paper argues that the PNC Combination is the optimal combining Confidence Distribution to solve the Behrens-Fisher Problem.

References

- [1] Xu, J.Q. (2011) The Generalized Confidence Interval of the Behrens-Fisher Problem. *Statistics and Decision*, **2**, 29-30.
- [2] Xie, M.G. and Singh, K. (2013) Confidence Distribution, the Frequentist Distribution Estimator of a Parameter: A Review. *International Statistical Review*, **81**, 3-39.
- [3] Schweder, T. and Hjort, N.L. (2002) Confidence and Likelihood. *Scandinavian Journal of Statistics*, **29**, 309-332. <http://dx.doi.org/10.1111/1467-9469.00285>
- [4] Singh, K., Xie, M. and Strawderman, W.E. (2005) Combining Information from Independent Sources through Confidence Distributions. *Annals of Statistics*, **33**, 159-183. <http://dx.doi.org/10.1214/009053604000001084>
- [5] Singh, K., Xie, M. and Strawderman, W.E. (2007) Confidence Distribution (CD)-Distribution Estimator of a Parameter. In Complex Datasets and Inverse Problems. *IMS Lecture Notes-Monograph Series*, **54**, 132-150. <http://dx.doi.org/10.1214/074921707000000102>
- [6] Tao, W.Y. and Ye, W.Z. (2016) The Confidence Distribution Method to the Behrens-Fisher Problem. *Journal of Applied Mathematics and Physics*, **4**, 286-293. <http://dx.doi.org/10.4236/jamp.2016.42036>
- [7] Singh, K., Xie, M. and Strawderman, W.E. (2005) Combining Information from Independent Sources through Confidence Distributions. *Annals of Statistics*, **33**, 159-183. <http://dx.doi.org/10.1214/009053604000001084>
- [8] Singh, K., Xie, M. and Strawderman, W. (2001) Confidence Distributions—Concept, Theory and Applications. Technical Report, Department of Statistics, Rutgers University, New Jersey.



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing a 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>