

# The Confidence Distribution Method to the Behrens-Fisher Problem

Wenyong Tao, Wanzhou Ye

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

Received 19 January 2016; accepted 22 February 2016; published 25 February 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

We use the methods of “The Welch-Satterthwaite test”, “The Cochran-Cox test”, “The Generalized  $p$ -value test”, “Computational Approach test” to structure different Confidence Distributions, and use the Confidence Distributions to give an new solution the confidence interval of the difference between two population means where the populations are assumed to be normal with unknown and unequal variances. Finally, we find the most effective solution through the numerical simulation.

## Keywords

Behrens-Fisher Problem, 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 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. A nice aspect of using a distribution estimator is that it contains a wealth of information for almost all types of inference. In frequentist inference, however, we often use a single point or an interval to estimate a parameter of interest. A simple question is: Can we also use a distribution function, or a “distribution estimator”, to estimate a parameter of interest in frequentist inference in the style of a Bayesian posterior?

Confidence Distribution is one such a “distribution estimator” that can be defined and interpreted in a frequentist framework, in which the parameter is a fixed and non-random quantity. The concept of confidence distribution has a long history, especially with its early interpretation associated with fiducial reasoning. Historically, it has been long misconstrued as a fiducial concept, and has not been fully developed under the frequentist framework. One nice aspect of treating a confidence distribution as a purely frequentist concept is that the confidence distribution is now a clean and coherent frequentist concept. Recently as the development of Confidence Distribution’s applications (such as Monte Carlo simulation based on Confidence Distribution), this approach has again aroused general concern.

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.

The main work of this article is using the Confidence Distribution method to solve the Behrens-Fisher problem. Firstly, we define the Confidence Distribution; then, we construct Confidence Distribution through some test method and prove these distributions to meet the definition; finally we find the optimal solution Confidence Distribution through numerical simulation.

## 2. Definition

The concept of “confidence” was first introduced by Neyman (1934, 1937) in his seminal papers on confidence intervals, where frequentist repetition properties for confidence were clarified. The earliest use of the terminology “confidence distribution” that we can find so far in a formal publication is Cox (1958). But for a long time nobody gives a complete and specific definition for Confidence Distribution.

The following definition is proposed and utilized in Schweder & Hjort (2002) [3] and Singh *et al.* (2005, 2007) [4] [5].

**Definition 2.1:** Given:  $\Theta$  is the parameter space of the unknown parameter of interest  $\theta$ ;  $X$  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  $\theta$ , if

- 1) For each given  $x \in X$ ,  $H_n(\cdot)$  is a cumulative distribution function on  $\Theta$ ;
- 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.

**Theorem 2.1:** If for each given  $x \in X$ ,  $H_n(\cdot)$  is a cumulative distribution function on  $\Theta$ , 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  $U [0,1]$ .

**Proof:** The cumulative distribution function  $H_n(\cdot)$  has an inverse function if  $H_n(\cdot)$  is continuous. Given  $H_n(\cdot) = \tau_n^{-1}(\cdot)$ , for any  $\alpha \in (0,1)$  and  $\theta \in \Theta$ , we can get that

$$\{x : H_n(\theta) \leq \alpha\} = \{x : \theta \leq \tau_n^{-1}(\alpha)\}$$

At  $\theta = \theta_0$ , we can get that

$$P(H_n(\theta_0) \leq \alpha) = P(\theta_0 \leq \tau_n^{-1}(\alpha)) = \alpha$$

Thus,  $H_n(\theta) = \tau_n^{-1}(\alpha)$  follows the uniform distribution  $U [0, 1]$ .

### 3. Structure and Proof

There is no fixed method to structure Confidence Distribution, we only need to construct to meet the definition of confidence distribution. Here we use some test method of Behrens-Fisher problem to structure Confidence Distribution.

#### 3.1. WS Distribution

Firstly, we use one of the most widely used method—“Welch–Satterthwaite test” (WST) [6] [7] to structure Confidence Distribution. The known conclusion is:

$$|X_1 - X_2| \pm t_{k,\alpha/2}^* \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where

$$s_i = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 / (n_i - 1), \quad t_{k,\alpha/2}^*$$

is the  $\alpha/2$  quantile of a Student's t distribution with k degrees of freedom.

On the basis of the conclusion above, we can easily get a probability distribution function (PDF) of  $\theta$ :

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

**Theorem 3.1:**  $f_{ws}(\theta)$ 's corresponding cumulative distribution function  $F_{ws}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

**Proof:** According to the probability distribution function (PDF) of the Student's t distribution:

$$f_t(\theta) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{[\theta - \mu]^2}{k}\right)^{-\frac{k+1}{2}}$$

where  $k$  is the degree of freedom,  $\mu$  is the non-central parameter.  $f_{ws}(\theta)$  is derived from  $t(k, \bar{X}_1 - \bar{X}_2)$  multiplied by a constant, so  $F_{ws}(\theta)$  is a cumulative distribution function, meet the condition 1) in **Definition 2.1**; according to **Theorem 2.1**, meet the condition 2) in **Definition 2.1**. So  $f_{ws}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

#### 3.2. CC Distribution

The most prominent test after the WST is the one proposed by Cochran and Cox (1950) [8]. The known conclusion is:

$$|X_1 - X_2| \pm \frac{\sum_{i=1}^2 t_i (s_i^2/n_i)}{\sum_{i=1}^2 (s_i^2/n_i)}$$

where

$$s_i = \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 / (n_i - 1), \quad t_i = t_{n_i-1,\alpha/2}$$

is the  $\alpha/2$  quantile of a Student's t distribution with  $n_i$  degrees of freedom.

On the basis of the conclusion above, we can easily get a probability distribution function (PDF) of  $\theta$  :

$$\begin{aligned}
 & f\left(\sum_{i=1}^2 t_i (s_i^2/n_i)\right) \\
 &= \int_{-\infty}^{+\infty} \frac{n_1}{s_1^2} \frac{\Gamma\left(\frac{n_1}{2}\right)}{\sqrt{(n_1\pi)}\Gamma\left(\frac{n_1-1}{2}\right)} \left(1 + \frac{\frac{n_1^2}{s_1^4}(\theta - y)^2}{n_1 - 1}\right)^{-\frac{n_1}{2}} \cdot \frac{n_2}{s_2^2} \frac{\Gamma\left(\frac{n_2}{2}\right)}{\sqrt{(n_2-1)\pi}\Gamma\left(\frac{n_2-1}{2}\right)} \left(1 + \frac{\frac{n_2^2}{s_2^4}y^2}{n_2 - 1}\right)^{-\frac{n_2}{2}} dy \\
 &= \frac{n_1 n_2 \Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)}{s_1^2 s_2^2 \sqrt{(n_1-1)(n_2-1)} \pi \Gamma\left(\frac{n_1-1}{2}\right) \Gamma\left(\frac{n_2-1}{2}\right)} \int_{-\infty}^{+\infty} \left(1 + \frac{\frac{n_1^2}{s_1^4}(\theta - y)^2}{n_1 - 1}\right)^{-\frac{n_1}{2}} \left(1 + \frac{\frac{n_2^2}{s_2^4}y^2}{n_2 - 1}\right)^{-\frac{n_2}{2}} dy \\
 & f_{cc}(\theta) = \sum_{i=1}^2 \binom{s_i^2}{n_i} \cdot \frac{n_1 n_2 \Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)}{s_1^2 s_2^2 \sqrt{(n_1-1)(n_2-1)} \pi \Gamma\left(\frac{n_1-1}{2}\right) \Gamma\left(\frac{n_2-1}{2}\right)} \\
 & \quad \times \int_{-\infty}^{+\infty} \left(1 + \frac{\frac{n_1^2}{s_1^4}(\theta - (\bar{X}_1 - \bar{X}_2) - y)^2}{n_1 - 1}\right)^{-\frac{n_1}{2}} \left(1 + \frac{\frac{n_2^2}{s_2^4}y^2}{n_2 - 1}\right)^{-\frac{n_2}{2}} dy
 \end{aligned}$$

**Theorem 3.2:**  $f_{cc}(\theta)$ 's corresponding cumulative distribution function  $F_{cc}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

**Proof:**  $f_{cc}(\theta)$  is composed of two t distribution obtained through convolution formula, so the  $F_{cc}(\theta)$  is a cumulative distribution function, meet the condition 1) in **Definition 2.1**; according to **Theorem 2.1**, meet the condition 2) in **Definition 2.1**. So  $f_{cc}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

### 3.3. GP Distribution

Taking advantage of the computational resources that are available today, the generalized p-value test (GPT) [9] uses a suitable pivot to come up with a simple test procedure. The known conclusion is:

$$|X_1 - X_2| \pm Z_{\alpha/2} \left\{ \sum_{i=2}^2 \frac{s_i^2}{V_i} \right\}^{1/2}$$

where  $Z_{\alpha/2}$  is the  $\alpha/2$  quantile of a standard normal distribution;  $V_i \sim \chi_{n_i-1}^2$ .

On the basis of the conclusion above, we can easily get a probability distribution function (PDF) of  $\theta$  :

1) The probability distribution function (PDF) of  $y_i = \frac{s_i^2}{V_i}$  :

$$p(y_i) = \left( -\frac{s_i^2}{y_i} \right) \left\{ \frac{\left(\frac{1}{2}\right)^{\frac{n_i}{2}}}{\Gamma\left(\frac{n_i}{2}\right)} \left(\frac{s_i^2}{y_i}\right)^{\frac{n_i}{2}-1} e^{-\frac{s_i^2}{2y_i}} \right\}$$

2) The probability distribution function (PDF) of  $z = \sum_{i=1}^2 \frac{s_i^2}{V_i}$ :

$$\begin{aligned} p(z) &= \int_{-\infty}^{+\infty} p_{y_1}(z - y_2) p_{y_2}(y_2) dy_2 \\ &= \frac{\left(\frac{1}{2}\right)^{\frac{n_1+n_2-2}{2}}}{\Gamma\left(\frac{n_1-1}{2}\right)\Gamma\left(\frac{n_2-1}{2}\right)} \int_0^{+\infty} \frac{s_1^2 s_2^2}{(z - y_2)^2 y_2^2} \left(\frac{s_1^2 s_2^2}{(z - y_2) y_2}\right)^{\frac{\pi}{2}-1} e^{-\left(\frac{s_1^2}{2(z-y_2)} + \frac{s_2^2}{2y_2}\right)} dy_2 \end{aligned}$$

3) The probability distribution function (PDF) of  $x = \sqrt{z}$ :

$$p(x) = 2x \frac{\left(\frac{1}{2}\right)^{\frac{n_1+n_2-2}{2}}}{\Gamma\left(\frac{n_1-1}{2}\right)\Gamma\left(\frac{n_2-1}{2}\right)} \int_0^{+\infty} \frac{s_1^2 s_2^2}{(x^2 - y_2)^2 y_2^2} \left(\frac{s_1^2 s_2^2}{(x^2 - y_2) y_2}\right)^{\frac{\pi}{2}-1} e^{-\left(\frac{s_1^2}{2(x^2-y_2)} + \frac{s_2^2}{2y_2}\right)} dy_2$$

4) The probability distribution function (PDF) of  $\theta = \mu_1 - \mu_2 = Z \cdot x$ :

$$f_{gp}(\theta) = \int_{-\infty}^{+\infty} \frac{1}{2\pi} \exp\left(-\frac{\left(\frac{\theta - (\bar{X}_1 - \bar{X}_2)}{x}\right)^2}{2}\right) p(x) \frac{1}{|x|} dx$$

**Theorem 3.3:**  $f_{gp}(\theta)$ 's corresponding cumulative distribution function  $F_{gp}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

**Proof:**  $f_{gp}(\theta)$  is composed of two chi-square distribution and a normal distribution obtained through convolution formula, so the  $F_{gp}(\theta)$  is a cumulative distribution function, meet the condition 1) in **Definition 2.1**; according to **Theorem 2.1**, meet the condition 2) in **Definition 2.1**. So  $f_{gp}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

### 3.4. CA Distribution

According to the classical Likelihood Ratio Tests and parameters bootstrap method, Chang and Pal (2008) [10] proposed Computational Approach Test (CAT). The known conclusion is:

$$|X_1 - X_2| \pm Z_{\alpha/2} \left( \hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2 \right)^{1/2}$$

where  $Z_{\alpha/2}$  is the  $\alpha/2$  quantile of a Standard normal distribution; parameters in  $\hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2$  are the solutions of the following equations:

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

On the basis of the conclusion above, we can easily get a probability distribution function (PDF) of  $\theta$ :

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

**Theorem 3.4:**  $f_{ca}(\theta)$ 's corresponding cumulative distribution function  $F_{ca}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

**Proof:**  $f_{ca}(\theta)$  is a normal distribution with a mean value of  $(\bar{X}_1 - \bar{X}_2)$  and a variance of  $\left[ \hat{\sigma}_{1(RML)}^2/n_1 + \hat{\sigma}_{2(RML)}^2/n_2 \right]$ , so the  $F_{ca}(\theta)$  is a cumulative distribution function, meet the condition 1) in **Definition 2.1**; according to **Theorem 2.1**, meet the condition 2) in **Definition 2.1**. So  $f_{ca}(\theta)$  is a Confidence Distribution of  $\theta = \mu_1 - \mu_2$ .

In this section, we construct four different Confidence Distributions to solve the Behrens-Fisher problem. Through this method, we the Confidence Distribution method, we can get almost all of the information of the parameter  $\theta = \mu_1 - \mu_2$ .

## 4. Numerical Simulation

### 4.1. 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:

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

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

$$\sum_{i=1}^n \frac{I(F_{\alpha/2} < \bar{X}_1^i - \bar{X}_2^i < F_{1-\alpha/2})}{n}$$

where,  $I$  is a indicative function. The more  $\hat{\eta}$  is close to  $\alpha$ , 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.

### 4.2. Optimality

Both  $H_1(\cdot)$  and  $H_2(\cdot)$  are Confidence Distribution of  $\theta$ , 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] [11].

### 4.3. 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 (**Table 1**).

According to the result of numerical simulation, we can see:

- 1) With the increase of sample size, the effectiveness of each Confidence Distribution increase;
- 2) In the small sample size, the effectiveness of ws and ca is relatively high;
- 3) In the relatively big sample size, cc, ws, ca are relatively stable and highly effective.

## 5. Conclusion

We construct four different Confidence Distributions. Through the numerical simulation we can find the optimal Confidence Distribution. In small sample size and relatively big sample size, the effectiveness of ws care rela-

**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 $\sigma_1^2/\sigma_2^2$ |         |         |         |         |         |
|--------------|-------------------------|--|---------|---------|---------|---------|---------|
|              |                         | 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 |
|              | cc                      | 0.03434  | 0.03055 | 0.03165 | 0.03644 | 0.04346 | 0.04502 |
|              | gp                      | 0.02250  | 0.02140 | 0.02230 | 0.02560 | 0.02880 | 0.03160 |
|              | ca                      | 0.04623  | 0.04372 | 0.04188 | 0.03812 | 0.03595 | 0.02665 |
| (10,10)      | ws                      | 0.05188  | 0.04839 | 0.04753 | 0.04914 | 0.05024 | 0.05241 |
|              | cc                      | 0.04484  | 0.03752 | 0.03537 | 0.03864 | 0.04152 | 0.04535 |
|              | gp                      | 0.03520  | 0.02930 | 0.02880 | 0.03020 | 0.03320 | 0.03500 |
|              | ca                      | 0.04256  | 0.04820 | 0.04873 | 0.04724 | 0.04582 | 0.04413 |
| (10,25)      | ws                      | 0.05027  | 0.04988 | 0.04931 | 0.04601 | 0.04079 | 0.03856 |
|              | cc                      | 0.04852  | 0.04565 | 0.04140 | 0.04126 | 0.04155 | 0.04287 |
|              | gp                      | 0.03910  | 0.03380 | 0.03600 | 0.03820 | 0.03850 | 0.04150 |
|              | ca                      | 0.04182  | 0.04590 | 0.04836 | 0.04993 | 0.04918 | 0.04825 |
| (25,25)      | ws                      | 0.04943  | 0.04893 | 0.04977 | 0.04971 | 0.05103 | 0.05054 |
|              | cc                      | 0.04827  | 0.04434 | 0.04498 | 0.04455 | 0.04619 | 0.04806 |
|              | gp                      | 0.04230  | 0.04050 | 0.03870 | 0.04120 | 0.04250 | 0.04370 |
|              | ca                      | 0.04927  | 0.04982 | 0.04979 | 0.04930 | 0.04791 | 0.04632 |
| (25,50)      | ws                      | 0.05004  | 0.05028 | 0.04868 | 0.04860 | 0.04694 | 0.04606 |
|              | cc                      | 0.04863  | 0.04744 | 0.04626 | 0.04471 | 0.04649 | 0.04848 |
|              | gp                      | 0.04530  | 0.04600 | 0.04470 | 0.04250 | 0.04380 | 0.04520 |
|              | ca                      | 0.04837  | 0.04962 | 0.04938 | 0.04987 | 0.04880 | 0.04973 |
| (50,50)      | ws                      | 0.04986  | 0.05027 | 0.04988 | 0.04900 | 0.04970 | 0.04990 |
|              | cc                      | 0.04876  | 0.04757 | 0.04747 | 0.04731 | 0.04894 | 0.04998 |
|              | gp                      | 0.04930  | 0.04900 | 0.04760 | 0.04500 | 0.04680 | 0.04820 |
|              | ca                      | 0.04928  | 0.04904 | 0.04978 | 0.04885 | 0.04903 | 0.04820 |

tively close, so we can compare the length of the confidence interval. Due tows come from the Student's t distribution, ca comes from the standard normal distribution and the Student's t distribution is a fat tail distribution, so ca is better than cc and ws. Therefore, this paper argues that the ca distribution is the optimal 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. *Complex Datasets and Inverse Problems*, **54**, 132-150. <http://dx.doi.org/10.1214/074921707000000102>
- [6] Welch, B.L. (1949) Further Notes on Mrs. Aspin's Tables. *Biometrika*, **36**, 293-296.
- [7] Satterthwaite, F.E. (1946) An Approximate Distribution of Estimates of Variance Components. *Biometrics Bulletin*, **2**, 110-114. <http://dx.doi.org/10.2307/3002019>
- [8] Cochran, W.G. and Cox, G.M. (1950) *Experimental Designs*. John Wiley and Sons, New York.
- [9] Weerahandi, S. (1994) *Exact Statistical Methods for Data Analysis (174-181)*. Springer Series in Statistics, Springer-Verlag, New York.
- [10] Chang, C.H. and Pal, N. (2008) A Revisit to the Behrens-Fisher Problem: Comparison of Five Test Methods. *Communications in Statistics—Simulation and Computation*, **37**, 1064-1085.
- [11] Singh, K., Xie, M. and Strawderman, W. (2001) *Confidence Distributions—Concept, Theory and Applications*. Technical Report, Dept. of Statistics, Rutgers University, New Jersey, USA.