Japanese Journal of Biometrics Vol. 38, No. 1, 1–16 (2017)

Original Article

Bayesian Indexes of Superiority and Equivalence and the *p*-value of the *F*-test for the Variances of Normal Distributions

Masaaki Doi[*1,*2], Kazuki Ide[*3,*4,*5] and Yohei Kawasaki[*3,*5]

[*1]Toray Industries, Inc. Clinical Data Science & Quality Management Department
[*2]Graduate School of Science and Engineering, Chuo University
[*3]Department of Pharmacoepidemiology, Graduate School of Medicine and Public Health, Kyoto University
[*4]Center for the Promotion of Interdisciplinary Education and Research, Kyoto University
[*5]Department of Drug Evaluation & Informatics, School of Pharmaceutical Sciences, University of Shizuoka
e-mail:Masaaki_Doi@nts.toray.co.jp

We here consider the problem of comparing the variances of two normal populations. To make a more efficient decision than that made with the conventional F-test, we propose using the Bayesian index of the superiority of the variance of one group to the other $\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$. We express this index according to the hypergeometric series and the cumulative distribution functions of well-known distributions. Furthermore, we investigate the relationship between the Bayesian index and the p-value of the F-test. In addition, we propose another index, the Bayesian index of equivalence of two groups, $\theta_e(\Delta) = Pr(\Delta < \sigma_1/\sigma_2 < 1/\Delta \mid \mathbf{x}_1, \mathbf{x}_2)$ for $0 < \Delta < 1$, which is also expressed according to the hypergeometric series and the cumulative distribution functions of well-known distributions. Finally, we evaluate the properties of the Bayesian index of equivalence using simulation, and illustrate the application of the Bayesian indexes with data from actual clinical trials.

Key words: Bayesian index; Conditional power prior; F-test; Historical data; Hypergeometric series; Variance of normal distribution.

1. Introduction

In biomedical studies, we encounter occasions to compare the variances in variables of interest across different conditions. Such occasions may be divided into two different situations. The first situation is when we mainly focus on comparing the variances. For example, test-retest variabilities (TRV) of visual acuity measurements are compared across different degrees of optical defocus in Rosser et al. (2004) and across different methods of scoring in Bosch and Wall (1997). The second situation is when we mainly focus on the location parameters (e.g., mean), and we

Received October 2016. Revised January 2017. Accepted February 2017.

want to check the assumption about the variances in the statistical method for comparing them. In many clinical trials with continuous outcomes, linear (mixed) models including t-test, ANOVA, and ANCOVA are used as the method of the primary analysis. Based on whether the variances are equal or not, we may change the statistical method (e.g., Student's t-test or Welch's t-test) because inappropriate choice of the method may lead to incorrect conclusions. See, for example, Welch (1938) and Glass et al. (1972). Therefore, to choose the correct method is important. However, for many clinical trials in this situation, the tests comparing the variances are known to have lower power than expected. See, for example, Markowski and Markowski (1990) and Wilcox (1995). This may occur because the sample sizes are calculated for comparing location parameters, which reduces the power of the test for comparing variances.

For a general two-group comparison of parameters, Bayesian approaches have gained increasing attention for their potential superiority in decision making compared to conventional frequentist methods, because a Bayesian approach can borrow strength from the historical data. For example, with a binomial distribution $B(n_i, p_i)$, Altham (1969), Kawasaki and Miyaoka (2012b), Zaslavsky (2013), and Kawasaki et al. (2014) considered the posterior probability $Pr(p_1 > p_1 \mid X_1, X_2)$. For the Poisson distribution $Po(\lambda_i)$, Kawasaki and Miyaoka (2012a) and Doi (2016) considered $Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$. Kawasaki and Miyaoka (2012b) referred to these types of probabilities as Bayesian indexes. For both distributions, the Bayesian indexes were shown to be expressed by the hypergeometric series, and the relationship between the Bayesian indexes and the p-values of conventional frequentist tests were investigated.

In this paper, we consider the problem of comparing the variances of two normal populations. F-test is most frequently used in this situation. To achieve a more effective decision than possible with the F-test by borrowing strength from the historical data, we propose a Bayesian index of superiority and equivalence for comparing the variances of two groups of normally distributed data.

The remainder of this paper is structured as follows. In section 2, we propose the Bayesian index of superiority for three situations, express these indexes by the hypergeometric series and the cumulative distribution functions of well-known distributions, and investigate their relationship with the p-values of the F-test. In section 3, we propose the Bayesian index of equivalence, which is also expressed by the hypergeometric series and the cumulative distributions functions. In section 4, we present the results of a Monte Carlo simulation to investigate the properties of $\theta_e(\Delta) \geq \gamma$ for several Δ and γ values used in the Bayesian index of equivalence. In section 5, we apply the Bayesian indexes to analyses of real data from actual clinical trials. Finally, we offer concluding remarks and highlight the prospects of these indexes in section 6.

2. Bayesian Index of Superiority

2.1 Definition and the Fundamental Theorem

For $i = 1, 2$, and $n_1, n_2 \in \mathbb{N}$, let $X_{i1}, \ldots, X_{in_i}$ be independent normal random variables with mean μ_i and variance σ_i^2. Let the realized values of $X_{i1}, \cdots, X_{in_i}$ be denoted by $\mathbf{x}_i = (x_{i1}, \ldots, x_{in_i})'$. For Bayesian analysis, let the prior distribution of σ_i^2 be the scaled inverse χ^2 distribution *Scaled-inv-*$\chi^2(\nu_i, \tau_i^2)$ for $\nu_i, \tau_i^2 > 0$, whose probability density function is

$$f(\sigma_i^2 \mid \nu_i, \tau_i^2) = \frac{\left(\nu_i \tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma\left(\nu_i/2\right)} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right).$$

This is equivalent to the inverse gamma distribution $Inv\text{-}Ga\left(\nu_i/2, \nu_i \tau_i^2/2\right)$. To compare the variances of two groups, we propose the Bayesian index of superiority as follows:

$$\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2).$$

In the following description, we first consider the case where μ_1 and μ_2 are known, and next consider the case where both means are unknown. In each case, the following theorem is crucial.

Theorem 1 If the (marginal) posterior distribution of σ_i^2 is $Inv\text{-}Ga(a_i, b_i)$ for $i = 1, 2$, then the Bayesian index $\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ has the following three expressions:

$$\begin{aligned}\theta =& 1 - \frac{1}{a_2 B(a_1, a_2)} \left(\frac{b_2}{b_1 + b_2}\right)^{a_2} \cdot {}_2F_1\left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 + b_2}\right)\\ =& I_{\frac{b_1}{b_1+b_2}}(a_1, a_2)\\ =& F_{2a_1, 2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\right),\end{aligned}$$

where

$${}_2F_1(a, b; c; z) = \sum_{t=0}^{\infty} \frac{(a)_t (b)_t}{(c)_t} \cdot \frac{z^t}{t!} \qquad (|z| < 1)$$

is the hypergeometric series, and $(k)_t = k(k+1)\cdots(k+t-1)$ for $t \in \mathbb{N}$ and $(k)_0 = 1$ is the Pochhammer symbol,

$$F_{\nu_1, \nu_2}(x) = \int_0^x \frac{1}{zB\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \left(\frac{\nu_1 z}{\nu_1 z + \nu_2}\right)^{\frac{\nu_1}{2}} \left(\frac{\nu_2}{\nu_1 z + \nu_2}\right)^{\frac{\nu_2}{2}} dz$$

is the cumulative distribution function of the F distribution $F(\nu_1, \nu_2)$, and

$$I_x(a, b) = \frac{1}{B(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$$

is the cumulative distribution function of the beta distribution $Beta(a, b)$, also known as the regularized incomplete beta function.

proof Let $\lambda_i = 1/\sigma_i^2$ be the precision, then

$$\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$$

$$= Pr(\lambda_1 < \lambda_2 \mid \mathbf{x}_1, \mathbf{x}_2). \tag{1}$$

When the (marginal) posterior distribution of σ_i^2 is $\mathit{Inv\text{-}Ga}(a_i, b_i)$, the (marginal) posterior distribution of λ_i is $Ga(a_i, b_i)$, whose probability density function is $f(\lambda_i \mid a_i, b_i) = b_i^{a_i}/\Gamma(a_i) \cdot \lambda_i^{a_i-1} \exp(-b_i\lambda_i)$. Hence, (1) is the Bayesian index for the Poisson parameters defined in Kawasaki and Miyaoka (2012a). Therefore, theorem 1 follows from Kawasaki and Miyaoka (2012a) and theorem 1 of Doi (2016). □

From the cumulative distribution function expressions in theorem 1, θ can be quite easily calculated using standard statistical software.

Remark 1 For theorem 1, since only the (marginal) posterior distribution of σ_i^2 is supposed as the inverse gamma distribution, the prior distribution of σ_i^2 may be improper as long as the posterior is the inverse gamma.

2.2 Case 1: μ_1 and μ_2 are Known

2.2.1 Calculation of the Bayesian Index of Superiority

In this case, we denote the likelihood of σ_i^2 by $L(\sigma_i^2 \mid \mathbf{x}_i, \mu_i)$. Since we suppose that the prior distribution of σ_i^2 is $\mathit{Scaled\text{-}inv\text{-}}\chi^2(\nu_i, \tau_i^2)$, the posterior distribution of σ_i^2 can be derived as follows:

$$\begin{aligned}
& f(\sigma_i^2 \mid \mathbf{x}_i, \mu_i, \nu_i, \tau_i^2) \\
\propto & L(\sigma_i^2 \mid \mathbf{x}_i, \mu_i) \cdot f(\sigma_i^2 \mid \nu_i, \tau_i^2) \\
= & \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij} - \mu_i)^2}{2\sigma_i^2}\right) \cdot \frac{\left(\nu_i\tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto & \left(\sigma_i^2\right)^{-(\nu_i+n_i)/2-1} \exp\left(-\frac{\nu_i\tau_i^2 + n_i \cdot T_i^2}{2\sigma_i^2}\right),
\end{aligned}$$

where

$$T_i^2 = \frac{1}{n_i} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2.$$

Hence, the posterior distribution of σ_i^2 is $\mathit{Inv\text{-}Ga}((\nu_i + n_i)/2,\ (\nu_i\tau_i^2 + n_i \cdot T_i^2)/2)$. Therefore, from theorem 1, we have

$$\theta = F_{\nu_1+n_1, \nu_2+n_2}\left(\frac{(\nu_1\tau_1^2 + n_1 \cdot T_1^2)/(\nu_1 + n_1)}{(\nu_2\tau_2^2 + n_2 \cdot T_2^2)/(\nu_2 + n_2)}\right). \tag{2}$$

2.2.2 The Relationship between the Bayesian Index of Superiority and the p-value of the one-sided F-test

Here, we consider the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$ when μ_1 and μ_2 are known. Under H_0, the test statistics T_1^2/T_2^2 follow $F(n_1, n_2)$. Hence, the p-value is calculated as

$$p = \int_{T_1^2/T_2^2}^{\infty} \frac{1}{zB\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \left(\frac{n_1 z}{n_1 z + n_2}\right)^{\frac{n_1}{2}} \left(\frac{n_2}{n_1 z + n_2}\right)^{\frac{n_2}{2}} dz$$

$$= 1 - F_{n_1,n_2}(T_1^2/T_2^2). \tag{3}$$

Then, the following theorem holds.

Theorem 2 If μ_1 and μ_2 are known and the prior distribution of σ_i^2 is *Scaled-inv-*$\chi^2(\nu_i, \tau_i^2)$ for $i = 1, 2$, then the following relation holds between the Bayesian index $\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \theta = 1 - p.$$

proof From (2) and (3),

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \theta = F_{n_1,n_2}(T_1^2/T_2^2) = 1 - p$$

holds. □

Remark 2 For the prior distribution,

$$\begin{aligned} f(\sigma_i^2 \mid \nu_i, \tau_i^2) &= \frac{\left(\nu_i \tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right) \\ &\propto (\sigma_i^2)^{-\nu_i/2-1} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right) \\ &\xrightarrow[\nu_i \to 0]{} (\sigma_i^2)^{-1}, \end{aligned}$$

when the prior distribution of σ_i^2 is $f(\sigma_i^2) \propto (\sigma_i^2)^{-1}$, which is improper, the posterior distribution of σ_i^2 is *Inv-Ga*$(n_i/2, n_i \cdot T_i^2/2)$. Therefore, as stated in remark 1, the Bayesian index can be expressed by theorem 1 as

$$\theta = F_{n_1,n_2}\left(T_1^2/T_2^2\right).$$

Then, the following theorem holds.

Theorem 3 If μ_1 and μ_2 are known and the prior distribution of σ_i^2 is $f(\sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then

$$\theta = 1 - p$$

holds.

Since ν_i is the prior effective sample size of *Scaled-inv-*$\chi^2(\nu_i, \tau_i^2)$ as defined in Morita et al. (2008), theorem 3 can be interpreted as follows: the Bayesian index with prior effective sample size 0 for both groups is equal to $(1-p)$ of the one-sided F-test. Furthermore, with the prior *Scaled-inv-*$\chi^2(\nu_i, \tau_i^2)$, the Bayesian index can be interpreted as equal to the F-test with the prior information of α_i additional samples.

2.3 Case 2: μ_1 and μ_2 are Unknown

In this case, we consider two types of the prior distributions of (μ_i, σ_i^2). In each type, we denote the likelihood of (μ_i, σ_i^2) by $L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i)$. Furthermore, in the following, we denote the probability density function of the normal inverse gamma distribution $NIG(\mu_0, k, \alpha, \beta)$ for $\mu_0 \in \mathbb{R}, k, \alpha, \beta > 0$ by

$$
\begin{aligned}
& f(\mu, \sigma^2 \mid \mu_0, k, \alpha, \beta) \\
=& \sqrt{\frac{k}{2\pi\sigma^2}} \exp\left(-\frac{k(\mu-\mu_0)^2}{2\sigma^2}\right) \times \frac{\beta^\alpha}{\Gamma(\alpha)} (\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right).
\end{aligned}
$$

When (μ, σ^2) follows $NIG(\mu_0, k, \alpha, \beta)$, the marginal distribution of σ^2 is $\mathit{Inv\text{-}Ga}(\alpha, \beta)$.

2.3.1 Calculation of the Bayesian index of superiority for the scaled inverse χ^2 variance prior

We first suppose that the prior distribution of μ_i is non-informative, i.e., $f(\mu_i) \propto 1$, and the prior distribution of σ_i^2 is $\mathit{Scaled\text{-}inv\text{-}}\chi^2(\nu_i, \tau_i^2)$. Then, the prior distribution of (μ_i, σ_i^2) is $\mathit{Scaled\text{-}inv\text{-}}\chi^2(\nu_i, \tau_i^2)$. Here, the posterior distribution of (μ_i, σ_i^2) can be derived as

$$
\begin{aligned}
& f(\mu_i, \sigma_i^2 \mid \mathbf{x}_i, \nu_i, \tau_i^2) \\
\propto& L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i) \cdot f(\mu_i, \sigma_i^2 \mid \nu_i, \tau_i^2) \\
=& \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij}-\mu_i)^2}{2\sigma_i^2}\right) \cdot \frac{\left(\nu_i\tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma\left(\nu_i/2\right)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto& \left(\sigma_i^2\right)^{-1/2} \exp\left(-\frac{n_i(\mu_i-\bar{x}_i)^2}{2\sigma_i^2}\right) \cdot \left(\sigma_i^2\right)^{-\frac{\nu_i+n_i-1}{2}-1} \exp\left(-\frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2}{2\sigma_i^2}\right),
\end{aligned}
$$

where

$$
\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}, S_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2.
$$

Therefore, the posterior distribution of (μ_i, σ_i^2) is

$$
NIG\left(\bar{x}_i, n_i,\ (\nu_i+n_i-1)/2,\ (\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2)/2\right),
$$

and the marginal posterior distribution of σ_i^2 is

$$
\mathit{Inv\text{-}Ga}((\nu_i+n_i-1)/2,\ (\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2)/2).
$$

Then, the Bayesian index can be expressed by theorem 1 as

$$
\theta = F_{\nu_1+n_1-1, \nu_2+n_2-1}\left(\frac{(\nu_1\tau_1^2 + (n_1-1)\cdot S_1^2)/(\nu_1+n_1-1)}{(\nu_2\tau_2^2 + (n_2-1)\cdot S_2^2)/(\nu_2+n_2-1)}\right). \tag{4}
$$

2.3.2 The relationship between the Bayesian index of superiority for the scaled inverse χ^2 variance prior and the p-value of the one-sided F-test

Here, we consider the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$ when μ_1 and μ_2 are unknown. Under H_0, the test statistics S_1^2/S_2^2 follow $F(n_1-1, n_2-1)$. Therefore, the p-value is calculated as

$$\begin{aligned} p =& \int_{S_1^2/S_2^2}^{\infty} \frac{1}{zB\left(\frac{n_1-1}{2}, \frac{n_2-1}{2}\right)} \\ &\times \left(\frac{(n_1-1)z}{(n_1-1)z+(n_2-1)}\right)^{\frac{n_1-1}{2}} \left(\frac{n_2-1}{(n_1-1)z+(n_2-1)}\right)^{\frac{n_2-1}{2}} dz \\ =& 1 - F_{n_1-1,n_2-1}(S_1^2/S_2^2). \end{aligned} \tag{5}$$

Then, the following theorem holds.

Theorem 4 If the prior distribution of μ_i is non-informative, i.e., $f(\mu_i) \propto 1$, and that of σ_i^2 is *Scaled-inv*-$\chi^2(\nu_i, \tau_i^2)$ for $i = 1, 2$, respectively, then the following relation holds between the Bayesian index $\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \theta = 1 - p.$$

proof From (4) and (5),

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \theta = F_{n_1-1,n_2-1}(S_1^2/S_2^2) = 1 - p$$

holds. □

Remark 3 For the prior distribution,

$$\begin{aligned} f(\mu_i, \sigma_i^2 \mid \nu_i, \tau_i^2) &= \frac{\left(\nu_i \tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right) \\ &\propto (\sigma_i^2)^{-\nu_i/2-1} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right) \\ &\xrightarrow[\nu_i \to 0]{} (\sigma_i^2)^{-1}, \end{aligned}$$

when the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$, the posterior distribution of (μ_i, σ_i^2) is $NIG(\bar{x}_i, n_i,\ (n_i-1)/2,\ (n_i-1) \cdot S_i^2/2)$ and the marginal posterior distribution of σ_i^2 is *Inv-Ga*$((n_i-1)/2,\ (n_i-1) \cdot S_i^2/2)$. Therefore, the Bayesian index can be expressed by theorem 1 as

$$\theta = F_{n_1-1,n_2-2}\left(S_1^2/S_2^2\right).$$

Then, the following theorem holds.

Theorem 5 If the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then

$$\theta = 1 - p$$

holds.

2.3.3 Calculation of the Bayesian index of superiority for the normal inverse gamma prior

We next suppose that the prior distribution is $\mu_i \mid \sigma_i^2 \sim N(\mu_{0,i}, \sigma_i^2/k_i)$ and $\sigma_i^2 \sim \textit{Scaled-inv-}\chi^2(\nu_i, \tau_i^2) = \textit{Inv-Ga}(\nu_i/2, \nu_i\tau_i^2/2)$. Then, the prior distribution of (μ_i, σ_i^2) is the normal inverse gamma distribution $NIG\left(\mu_{0,i}, k_i, \nu_i/2, \nu_i\tau_i^2/2\right)$. Hence, the posterior distribution of (μ_i, σ_i^2) can be derived as

$$\begin{aligned}
&f(\mu_i, \sigma_i^2 \mid \mathbf{x}_i, \mu_{0,i}, k_i, \nu_i, \tau_i^2) \\
\propto &L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i) \cdot f(\mu_i, \sigma_i^2 \mid \mu_{0,i}, k_i, \nu_i, \tau_i^2) \\
= &\frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij}-\mu_i)^2}{2\sigma_i^2}\right) \cdot \left(\frac{k_i}{2\pi\sigma_i^2}\right)^{1/2} \exp\left(-\frac{k_i(\mu_i-\mu_{0,i})^2}{2\sigma_i^2}\right) \\
&\times \frac{\left(\nu_i\tau_i^2/2\right)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma\left(\nu_i/2\right)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto &\left(\sigma_i^2\right)^{-1/2} \exp\left(-\frac{(k_i+n_i)(\mu_i-\mu_{ni})^2}{2\sigma_i^2}\right) \\
&\times \left(\sigma_i^2\right)^{-\frac{\nu_i+n_i}{2}-1} \cdot \exp\left(-\frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2 + \dfrac{k_i n_i(\mu_{0,i}-\bar{x}_i)^2}{k_i+n_i}}{2\sigma_i^2}\right),
\end{aligned}$$

where

$$\begin{aligned}
\mu_{ni} &= \frac{k_i\mu_{0,i} + n_i\bar{x}_i}{k_i+n_i},\ k_{ni} = k_i + n_i, \\
a_{ni} &= \frac{\nu_i+n_i}{2},\ b_{ni} = \frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2}{2} + \frac{k_i n_i(\mu_{0,i}-\bar{x}_i)^2}{2(k_i+n_i)}.
\end{aligned}$$

Then, the posterior distribution of (μ_i, σ_i^2) is $NIG(\mu_{ni}, k_{ni}, a_{ni}, b_{ni})$. Hence, the marginal posterior distribution of σ_i^2 is $\textit{Inv-Ga}(a_{ni}, b_{ni})$. Therefore, the Bayesian index can be expressed from theorem 1 as

$$\theta = F_{2a_{n1}, 2a_{n2}}\left(\frac{\dfrac{\nu_1\tau_1^2 + (n_1-1)\cdot S_1^2}{\nu_1+n_1} + \dfrac{k_1 n_1(\mu_{0,1}-\bar{x}_1)^2}{(\nu_1+n_1)(k_1+n_1)}}{\dfrac{\nu_2\tau_2^2 + (n_2-1)\cdot S_2^2}{\nu_2+n_2} + \dfrac{k_2 n_2(\mu_{0,2}-\bar{x}_2)^2}{(\nu_2+n_2)(k_2+n_2)}}\right). \tag{6}$$

2.3.4 The relationship between the Bayesian index of superiority for the normal inverse gamma prior and the p-value of the one-sided F-test

Then, the following theorem holds.

Theorem 6 If the prior distribution of (μ_i, σ_i^2) is $NIG(\mu_{0,i}, k_i, \nu_i/2, \nu_i\tau_i^2/2)$ for $i = 1, 2$, then the following relation holds between the Bayesian index $\theta = Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1, k_1, \tau_1^2, \nu_2, k_2, \tau_2^2) \to (-1, 0, 0, -1, 0, 0)} \theta = 1 - p.$$

proof From (6) and (5),

$$\lim_{(\nu_1,k_1,\tau_1^2,\nu_2,k_2,\tau_2^2)\to(-1,0,0,-1,0,0)} \theta = F_{n_1-1,n_2-1}(S_1^2/S_2^2)$$
$$= 1-p$$

holds. □

Remark 4 For the prior distribution,

$$\begin{aligned}
&f(\mu_i,\sigma_i^2 \mid \mu_{0,i},k_i,\nu_i,\tau_i^2)\\
&= \sqrt{\frac{k_i}{2\pi\sigma_i^2}}\exp\left(-\frac{k_i(\mu_i-\mu_{0,i})^2}{2\sigma_i^2}\right)\cdot\frac{\left(\nu_i\tau_i^2/2\right)^{\nu_i/2}}{\Gamma(\nu_i/2)}(\sigma_i^2)^{-\nu_i/2-1}\exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right)\\
&\propto (\sigma_i^2)^{-(\nu_i+1)/2-1}\exp\left(-\frac{\nu_i\tau_i^2+k_i(\mu_i-\mu_{0,i})^2}{2\sigma_i^2}\right)\\
&\xrightarrow[(\nu_i,\tau_i^2,k_i)\to(-1,0,0)]{} (\sigma_i^2)^{-1}.
\end{aligned}$$

As already shown in theorem 5, if the prior distribution of (μ_i,σ_i^2) is $f(\mu_i,\sigma_i^2)\propto(\sigma_i^2)^{-1}$, then

$$\theta = 1-p$$

holds.

2.4 Remark on the Prior Distribution

To utilize the historical data effectively, we here consider how to construct the prior distribution of (μ_i,σ_i^2). For $i=1,2$ and $j=1,\cdots,n_{0,i}$, let the historical data $x_{0,ij}$ independently follow $N(\mu_i,\sigma_i^2)$, and $\mathbf{x}_{0,i}=(x_{0,i1},\ldots,x_{0,in_{0,i}})'$, and let

$$f_0(\mu_i,\sigma_i^2)\propto\left(\sigma_i^2\right)^{-1}.$$

Here, for $0\le\alpha_i\le1$, an example of the conditional power prior distribution, defined in Ibrahim and Chen (2000), is

$$\begin{aligned}
f(\mu_i,\sigma_i^2) &\propto L(\mu_i,\sigma_i^2\mid\mathbf{x}_{0,i})^{\alpha_i}\cdot f_0(\mu_i,\sigma_i^2)\\
&\propto\left\{\frac{1}{(2\pi\sigma_i^2)^{n_{0,i}/2}}\exp\left(-\frac{1}{2\sigma_i^2}\sum_{j=1}^{n_{0,i}}(x_{0,ij}-\mu_i)^2\right)\right\}^{\alpha_i}\cdot\left(\sigma_i^2\right)^{-1}\\
&=\left(\sigma_i^2\right)^{-1/2}\exp\left(-\frac{\alpha_i n_{0,i}(\mu_i-\bar{x}_{0,i})^2}{2\sigma_i^2}\right)\\
&\quad\times\left(\sigma_i^2\right)^{-(\alpha_i n_{0,i}-1)/2-1}\exp\left(-\frac{\alpha_i(n_{0,i}-1)\cdot S_{0,i}^2}{2\sigma_i^2}\right),
\end{aligned}\tag{7}$$

where

$$\bar{x}_{0,i}=\frac{1}{n_{0,i}}\sum_{j=1}^{n_{0,i}}x_{0,ij},\ S_{0,i}^2=\frac{1}{n_{0,i}-1}\sum_{j=1}^{n_{0,i}}(x_{0,ij}-\bar{x}_{0,i})^2.$$

Then, the prior distribution of (μ_i, σ_i^2) is the normal inverse gamma distribution $NIG(\bar{x}_{0,i},\ \alpha_i n_{0,i},\ (\alpha_i n_{0,i}-1)/2,\ \alpha_i(n_{0,i}-1)\cdot S_{0,i}^2/2)$ when $\alpha_i n_{0,i} > 1$. Hence, the marginal prior distribution of σ_i^2 is $\mathit{Inv\text{-}Ga}((\alpha_i n_{0,i}-1)/2,\ \alpha_i(n_{0,i}-1)\cdot S_{0,i}^2/2)$ when $\alpha_i n_{0,i} > 1$. In this situation, the next corollary directly follows from theorem 6.

Corollary 1 If the prior distribution of (μ_i, σ_i^2) is the conditional power prior described above for $i = 1, 2$, then

$$\lim_{(\alpha_1,\alpha_2)\to(0,0)} \theta = 1 - p$$

holds.

3. Bayesian Index of Equivalence

Next, we propose the Bayesian index of equivalence for Δ satisfying $1 < \Delta$ as follows:

$$\theta_e(\Delta) = Pr(\,1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2)\,.$$

Here, we compare Δ and $1/\Delta$ not to the ratio of the variances but rather to the ratio of the standard deviations. Then, the following theorem holds.

Theorem 7 If the (marginal) posterior distribution of σ_i^2 is $\mathit{Scaled\text{-}inv\text{-}}\chi^2(a_i, b_i)$ for $i = 1, 2$, then

$$\begin{aligned}
\theta_e(\Delta) =& F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\Delta^2\right) - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\frac{1}{\Delta^2}\right)\\
=& I_{\frac{b_1\cdot\Delta^2}{b_1\cdot\Delta^2+b_2}}(a_1,a_2) - I_{\frac{b_1/\Delta^2}{b_1/\Delta^2+b_2}}(a_1,a_2)\\
=& \frac{1}{a_2 B(a_1,a_2)}\left(\frac{b_2}{b_1/\Delta^2+b_2}\right)^{a_2}\cdot {}_2F_1\left(a_2, 1-a_1; 1+a_2; \frac{b_2}{b_1/\Delta^2+b_2}\right)\\
&- \frac{1}{a_2 B(a_1,a_2)}\left(\frac{b_2}{b_1\cdot\Delta^2+b_2}\right)^{a_2}\cdot {}_2F_1\left(a_2, 1-a_1; 1+a_2; \frac{b_2}{b_1\cdot\Delta^2+b_2}\right).
\end{aligned}$$

proof Since $Pr(\sigma_1/\sigma_2 = \Delta) = 0$,

$$\begin{aligned}
\theta_e(\Delta) &= Pr(\,1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2)\\
&= Pr(\sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) - Pr(\sigma_1/\sigma_2 < 1/\Delta \mid \mathbf{x}_1, \mathbf{x}_2)\\
&= Pr(\sigma_1^2/\sigma_2^2 < \Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2) - Pr(\sigma_1^2/\sigma_2^2 < 1/\Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2).
\end{aligned}$$

Then, consider the posterior distribution of $\sigma_1^2/\sigma_2^2 = \lambda_2/\lambda_1$, where $\lambda_i = 1/\sigma_i^2$ is the precision for $i = 1, 2$. From theorem 3 in Doi (2016) or (2.10) in Price and Bonett (2000),

$$Pr(\lambda_2/\lambda_1 < c \mid \mathbf{x}_1, \mathbf{x}_2) = F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1}\cdot c\right),$$

therefore

$$\theta_e(\Delta) = Pr(\lambda_2/\lambda_1 < \Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2) - Pr(\lambda_2/\lambda_1 < 1/\Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2)$$

$$
\begin{aligned}
&= F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1}\cdot\Delta^2\right) - F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1}\cdot\frac{1}{\Delta^2}\right)\\
&= \left\{1 - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\frac{1}{\Delta^2}\right)\right\} - \left\{1 - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\Delta^2\right)\right\}\\
&\quad (\because F_{m,n}(1/x) = 1 - F_{n,m}(x))\\
&= F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\Delta^2\right) - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\cdot\frac{1}{\Delta^2}\right).
\end{aligned}
$$

The rest of the proof follows from theorem 3 in Doi (2016). □

4. Simulation

We conducted a Monte Carlo simulation to investigate the property of "$\theta_e(\Delta) \geq \gamma$" for several values of Δ and γ. We used the conditional power prior distribution. The historical data $x_{0,ij}$ independently follow $N(0, \sigma_i^2)$ for $i = 1, 2$; $j = 1, \ldots, n_{0,i}$, and we consider $\alpha_1 = \alpha_2 = 1$. The present data x_{ij} independently follow $N(0, \sigma_i^2)$ for $i = 1, 2$; $j = 1, \ldots, n_i$. Here, let $n = n_1 = n_2 = 25, 50, 100, 200$, and $n_0 = n_{0,1} = n_{0,2} = 0, 25, 50, 100, 200$, with $n_{0,i} \leq n_i$ for $i = 1, 2$. Further, we take $\gamma = 0.90, 0.95$ and $\Delta = 1.10, 1.25, 1.50, 2.00$. We conducted 100,000 iterations for each scenario. For the first scenario, we set $\sigma_1 = \sigma_2 = 10$; that is, the variances are equal. As shown in Table 1, the percentage satisfying this condition heavily depended on the sample size. For the second scenario, we set $\sigma_1 = 15$ and $\sigma_2 = 10$; that is, $\sigma_1/\sigma_2 = 1.5$, so that the variances of group 1 are greater than those of group 2. As shown in Table 2, the percentages satisfying $\theta_e(1.50) \geq 0.90$ and $\theta_e(1.50) \geq 0.95$ show minimal dependence on the sample size when $n \geq 50$. These results suggest that the decision of suitable values of Δ and γ must be considered depending on the

Table 1. The percentage satisfying $\theta_e(\Delta) \geq \gamma$ when $\sigma_1 = \sigma_2 = 10$.

n_0	n	$\gamma = 0.90$				$\gamma = 0.95$			
		$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0	25	0.00	0.00	48.03	95.92	0.00	0.00	0.00	90.87
25	25	0.00	0.00	87.31	99.94	0.00	0.00	76.07	99.82
0	50	0.00	0.00	87.53	99.95	0.00	0.00	75.81	99.81
25	50	0.00	45.08	97.06	100.00	0.00	0.00	93.19	100.00
50	50	0.00	64.46	99.31	100.00	0.00	41.04	98.16	100.00
0	100	0.00	64.56	99.37	100.00	0.00	49.84	98.23	100.00
25	100	0.00	76.76	99.88	100.00	0.00	59.48	99.57	100.00
50	100	0.00	84.82	99.97	100.00	0.00	71.66	99.89	100.00
100	100	0.00	93.67	100.00	100.00	0.00	86.49	100.00	100.00
0	200	0.00	93.70	100.00	100.00	0.00	86.48	100.00	100.00
25	200	0.00	95.86	100.00	100.00	0.00	90.93	100.00	100.00
50	200	0.00	97.50	100.00	100.00	0.00	93.95	100.00	100.00
100	200	5.09	98.93	100.00	100.00	0.00	97.26	100.00	100.00
200	200	43.87	99.84	100.00	100.00	0.00	99.46	100.00	100.00

Table 2. The percentage satisfying $\theta_e(\Delta) \geq \gamma$ when $\sigma_1 = 15$, and $\sigma_2 = 10$.

n_0	n	$\gamma = 0.90$				$\gamma = 0.95$			
		$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0	25	0.00	0.00	8.87	54.19	0.00	0.00	0.00	39.33
25	25	0.00	0.00	10.62	76.71	0.00	0.00	5.48	64.46
0	50	0.00	0.00	10.06	76.21	0.00	0.00	5.04	63.49
25	50	0.00	0.27	10.27	88.19	0.00	0.00	5.19	79.53
50	50	0.00	0.12	10.16	94.12	0.00	0.04	5.16	88.61
0	100	0.00	0.11	10.00	94.23	0.00	0.02	4.94	88.67
25	100	0.00	0.07	10.06	97.17	0.00	0.02	5.06	93.89
50	100	0.00	0.03	10.00	98.63	0.00	0.01	5.09	96.77
100	100	0.00	0.00	10.00	99.72	0.00	0.00	5.01	99.21
0	200	0.00	0.01	10.19	99.73	0.00	0.00	5.16	99.21
25	200	0.00	0.00	10.10	99.86	0.00	0.00	5.04	99.58
50	200	0.00	0.00	9.99	99.95	0.00	0.00	5.09	99.79
100	200	0.00	0.00	9.97	99.99	0.00	0.00	5.02	99.95
200	200	0.00	0.00	9.95	100.00	0.00	0.00	4.98	100.00

situation. For example, if $n = 50$ and $n_0 \geq 25$, then $\theta_e(1.50) \geq 0.90$ or $\theta_e(1.50) \geq 0.95$ may be appropriate.

5. Application

The application of the Bayesian indexes of superiority and equivalence was evaluated using data from actual clinical trials, as shown in Table 3. Trial (a) and (b) are two selected trials shown in Table 1 of Gould (1991). Here, we supposed that trial (a) is a previous trial and trial (b) is the present trial, and $i = 1, 2$ indicate the placebo and drug A group, respectively. Therefore, we utilized the data of trial (a) to specify the conditional power prior. We suppose that $\alpha = \alpha_1 = \alpha_2$, and take $\alpha = 0.0, 0.2, 0.5, 0.8, 1.0$. The prior distributions of (μ_i, σ_i^2) were derived from (7) with the following data

- $n_{0,1} = 47$, $\bar{x}_{0,1} = 3.04$, $S_{0,1}^2 = 9.20^2 \fallingdotseq 84.64$
- $n_{0,2} = 44$, $\bar{x}_{0,2} = 8.43$, $S_{0,2}^2 = 8.17^2 \fallingdotseq 66.75$,

and are shown in Table 4 for each α. Next, using the following data of trial (b)

- $n_1 = 53$, $\bar{x}_1 = 3.75$, $S_1^2 = 7.07^2 \fallingdotseq 49.98$
- $n_2 = 54$, $\bar{x}_2 = 10.20$, $S_2^2 = 9.39^2 \fallingdotseq 88.17$,

Table 3. Hypertention data

	Placebo ($i = 1$)			Drug A ($i = 2$)		
Trial	n	mean	SD	n	mean	SD
(a)	47	3.04	9.20	44	8.43	8.17
(b)	53	3.75	7.07	54	10.20	9.39

Table 4. Prior distributions of (μ_i, σ_i^2)

α	Placebo $(i=1)$	Drug A $(i=2)$
0.0	$f(\mu_1,\sigma_1^2) \propto (\sigma_1^2)^{-1}$	$f(\mu_2,\sigma_2^2) \propto (\sigma_2^2)^{-1}$
0.2	$NIG(3.04, 9.4, 4.2, 389.34)$	$NIG(8.43, 8.8, 3.9, 287.02)$
0.5	$NIG(3.04, 23.5, 11.3, 973.36)$	$NIG(8.43, 22.0, 10.5, 717.55)$
0.8	$NIG(3.04, 37.6, 18.3, 1557.38)$	$NIG(8.43, 35.2, 17.1, 1148.08)$
1.0	$NIG(3.04, 47.0, 23.0, 1946.72)$	$NIG(8.43, 44.0, 21.5, 1435.10)$

Table 5. Posterior distributions of (μ_i, σ_i^2).

α	Placebo $(i=1)$	Drug A $(i=2)$
0.0	$NIG(3.75, 53.0, 26.0, 1299.61)$	$NIG(10.20, 54.0, 26.5, 2336.56)$
0.2	$NIG(3.64, 62.4, 30.7, 1692.98)$	$NIG(9.95, 62.8, 30.9, 2647.29)$
0.5	$NIG(3.53, 76.5, 37.8, 2281.18)$	$NIG(9.69, 76.0, 37.5, 3103.08)$
0.8	$NIG(3.46, 90.6, 44.8, 2868.07)$	$NIG(9.50, 89.2, 44.1, 3551.40)$
1.0	$NIG(3.42, 100.0, 49.5, 3258.89)$	$NIG(9.41, 98.0, 48.5, 3847.62)$

Table 6. Marginal posterior distributions of σ_i^2 .

α	Placebo $(i=1)$	Drug A $(i=2)$
0.0	$Inv\text{-}Ga(26.0, 1299.61)$	$Inv\text{-}Ga(26.5, 2336.56)$
0.2	$Inv\text{-}Ga(30.7, 1692.98)$	$Inv\text{-}Ga(30.9, 2647.29)$
0.5	$Inv\text{-}Ga(37.8, 2281.18)$	$Inv\text{-}Ga(37.5, 3103.08)$
0.8	$Inv\text{-}Ga(44.8, 2868.07)$	$Inv\text{-}Ga(44.1, 3551.40)$
1.0	$Inv\text{-}Ga(49.5, 3258.89)$	$Inv\text{-}Ga(48.5, 3847.62)$

Table 7. Bayesian index of superiority and equivalence.

	Superiority		Equivalence			
α	θ	$1-\theta$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0.0	0.021	0.979	0.084	0.331	0.810	0.998
0.2	0.043	0.957	0.158	0.509	0.926	1.000
0.5	0.087	0.913	0.281	0.715	0.984	1.000
0.8	0.140	0.860	0.403	0.845	0.997	1.000
1.0	0.179	0.821	0.476	0.899	0.999	1.000

we derived the posterior distributions. The posterior distributions of (μ_i, σ_i^2) and the marginal posterior distributions of $\sigma_i^2$are shown in Table 5 and Table 6, respectively. Finally, the Bayesian indexes are shown in Table 7. For trial (a), the placebo group $(i=1)$ showed a larger standard deviation than the drug A group $(i=2)$. By contrast, for trial (b), the drug A group showed a larger standard deviation. According to the present data (trial (b)) only, that is, when $\alpha=0$, θ is quite small, which makes the variance of the placebo group seem greater. However, as α increases, i.e., the weight of the information of trial (a) increases, θ increases monotonically, and is no longer small. Furthermore, the p-value of the F-test with $H_0: \sigma_1^2=\sigma_2^2$ versus $H_1: \sigma_1^2>\sigma_2^2$ is 0.979, and, as shown in remark 4, is equal to $1-\theta$ with $\alpha=0.0$. Next, we consider the situation of $\theta_e(\Delta)$. Based on the result of the simulation in section 4., we assume that $\theta_e(1.50) \geq 0.95$ shows the equivalence, because the sample size is about 50 for both groups and for both the historical and present data. Then, when using only the present data ($\alpha = 0.0$) and $\alpha = 0.2$, the equivalence is not shown. By contrast, when $\alpha = 0.5, 0.8, 1.0$, that is, when the weight of the historical data is moderate to large, the equivalence is shown.

In order to apply these indexes to the real clinical trials, we have to consider whether α, Δ and γ can be pre-specified based on sufficiently reliable information. If we can pre-specify them suitably, we can determine the statistical method for comparing the means based on whether $\theta_e(\Delta) \geq \gamma$ holds or not. On the other hand, if we cannot pre-specify them, it may be hard to determine the statistical method for comparing the means based on whether $\theta_e(\Delta) \geq \gamma$ or not because it depends on the choice of α, Δ and γ. In such case, we have to determine the statistical method based only on the present trial data, and we can utilize $\theta_e(\Delta)$'s for several α's to scrutinize the appropriateness of the method. Depending on the values of $\theta_e(\Delta)$'s, we may conduct sensitivity analysis by changing the statistical method for comparing the means.

6. Conclusion

We have proposed the Bayesian index of superiority to make a more efficient decision for comparing the variances between two groups than possible with the conventional F-test. This index was expressed by the hypergeometric series and the cumulative distribution functions of well-known distributions. Furthermore, we showed that as the amount of prior information decreases, the Bayesian index of superiority approaches the $(1-p)$ value of the F-test with $H_0\colon \sigma_1^2 = \sigma_2^2$ versus $H_1\colon \sigma_1^2 > \sigma_2^2$. Moreover, if the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then $\theta = 1 - p$ holds. This indicates that the Bayesian index with a "non-informative" prior or "zero prior effective sample size" can have the same statistical properties as the F-test; however, with incorporation of suitable historical data, the Bayesian index can potentially be used to make a more efficient decision. In addition, we proposed the Bayesian index of equivalence $\theta_e(\Delta)$, which was evaluated with a Monte Carlo simulation. The results showed that the percentage satisfying $\theta_e(\Delta) \geq \gamma$ heavily depends on the sample size. Therefore, the appropriate values of Δ and γ must be decided on a case-by-case basis. If we mainly focus on comparing the variances, we can utilize the index of superiority and equivalence based on the objectives of trials. If we want to check the assumption about the variances in some statistical method, we can utilize the index of equivalence. In any case, in order to use these indexes for the confirmatory purpose, it is crucial to pre-specify $\alpha_1, \alpha_2, \Delta$, and γ suitably based on the sufficiently reliable information because θ and whether $\theta_e(\Delta) \geq \gamma$ or not depend on them. Therefore, the important future work is to develop a suitable method for constructing the prior distributions, including selecting suitable historical data, and deciding α_1, α_2 for the conditional power prior.

Acknowledgements

We would like to thank anonymous referees for their insightful comments.

REFERENCES

Altham, P. M. (1969). Exact Bayesian analysis of a 2×2 contingency table, and Fisher's "exact" significance test. *Journal of the Royal Statistical Society. Series B (Methodological)*, **31**(2): 261–269.

Bosch, M. E. V. and Wall, M. (1997). Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye*, **11**(3): 411–417.

Doi, M. (2016). Bayesian index of superiority and the p-value of the conditional test for Poisson parameters. *Journal of the Japan Statistical Society*, **46**(2): 99–127.

Glass, G. V., Peckham, P. D., and Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, **42**(3): 237–288.

Gould, A. L. (1991). Using prior findings to augment active-controlled trials and trials with small placebo groups. *Drug Information Journal*, **25**(3): 369–380.

Ibrahim, J. G. and Chen, M.-H. (2000). Power prior distributions for regression models. *Statistical Science*, **15**(1): 46–60.

Kawasaki, Y. and Miyaoka, E. (2012a). A Bayesian inference of $P(\lambda_1 < \lambda_2)$ for two Poisson parameters. *Journal of Applied Statistics*, **39**(10): 2141–2152.

Kawasaki, Y. and Miyaoka, E. (2012b). A Bayesian inference of $P(\pi_1 > \pi_2)$ for two proportions. *Journal of biopharmaceutical statistics*, **22**(3): 425–437.

Kawasaki, Y., Shimokawa, A., and Miyaoka, E. (2014). On the Bayesian index of superiority and the p-value of the Fisher exact test for binomial proportions. *Journal of the Japan Statistical Society*, **44**(1): 73–81.

Markowski, C. A. and Markowski, E. P. (1990). Conditions for the effectiveness of a preliminary test of variance. *The American Statistician*, **44**(4): 322–326.

Morita, S., Thall, P. F., and Müller, P. (2008). Determining the effective sample size of a parametric prior. *Biometrics*, **64**(2): 595–602.

Price, R. M. and Bonett, D. G. (2000). Estimating the ratio of two Poisson rates. *Computational Statistics & Data Analysis*, **34**(3): 345–356.

Rosser, D. A., Murdoch, I. E., and Cousens, S. N. (2004). The effect of optical defocus on the test–retest variability of visual acuity measurements. *Investigative ophthalmology & visual science*, **45**(4): 1076–1079.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, **29**(3/4): 350–362.

Wilcox, R. R. (1995). ANOVA: A paradigm for low power and misleading measures of effect size? *Review of Educational Research*, **65**(1): 51–77.

Zaslavsky, B. G. (2013). Bayesian hypothesis testing in two-arm trials with dichotomous outcomes. *Biometrics*, **69**(1): 157–163.