A Bayesian Approach to Assessing Population Bioequivalence in a 2×2 Crossover Design

Hyun Sook Oh [1] and Seoung-gon Ko [2]

Abstract

A Bayesian testing procedure is proposed for assessment of bioequivalence in both mean and variance which ensures population bioequivalence under normality assumption. We derive the joint posterior distribution of the means and variances in a standard 2×2 crossover experimental design and propose a Bayesian testing procedure for bioequivalence based on a Markov chain Monte Carlo methods. The proposed method is applied to a real data set.

Key Words : Bayesian hypothesis test, drug prescribability, Markov chain Monte Carlo.

1. Introduction

Population and individual bioequivalence have drawn considerable attention in bioequivalence study. Population bioequivalence (PBE) compares overall distributions of bioavailability while individual bioequivalence (IBE) compares bioavailabilities within individuals (Anderson and Hauck, 1990; Schall, 1995). In practice, PBE is considered for assessment of drug prescribability, i.e., whether a physician can prescribe a brand-name drug product or a generic drug product shown to be bioequivalent to the brand-name drug product for his/her new patient who has not previously received either of them. On the other hand, IBE refers to switchability, i.e., whether a patient who has been taking one formulation can switch to another with similar safety and efficacy.

Under the normality assumption, PBE can be established by demonstrating equivalence in both average and variability since a normal distribution is uniquely determined by its mean and variance. For assessment of bioequivalence in variance only, two testing procedures have been proposed by Liu and Chow(1992) and Wang(1997). Also, Chen et al.(1996) derived an exact confidence region approach for the assessment of bioequivalence in variance only when the intersubject variance is known, and considered a large sample approximation when the intersubject variance is unknown.

Simultaneous testing for the assessment of bioequivalence in average bioavailability and variability of bioavailability in a 2×2 crossover design has been considered by Hauck et al.(1997) in a frequentist framework. They derived a likelihood ratio procedure for a very simple model and modified the procedure for a more complicated model. However, their procedure still does not take into account of the full model.

[1]Department of Applied Statistics, Kyungwon University, Sungnam 461-701, Korea.
[2]Department of Applied Statistics, Kyungwon University, Sungnam 461-701, Korea.

A Bayesian approach to assessment of bioequivalence is given by Grieve (1985). However, only the average bioequivalence was considered in Grieve(1985) and simultaneous consideration of the average and variance was not handled.

In this paper we suggest a Bayesian testing procedure for assessing bioequivalence in mean and variance simultaneously under a conventional model on a standard 2×2 crossover design which is the most commonly used statistical design for comparing two drug formulations. We propose a model and establish interval hypotheses for simultaneous testing of bioequivalence in mean and variance. In the model, intra-subject variabilities are assumed to be the same from subject to subject but different from formulation to formulation. The proposed procedure takes into consideration of mean and variance simultaneously, and takes a Bayesian approach so that it can utilize useful prior information when available. Compared to Hauck et al.(1997), the proposed procedure is more general in that it can be applied to a more complex model on a 2×2 crossover design.

For Bayesian testing procedure we assume a noninformative prior for each unknown parameter in the model unless it brings posterior impropriety. When noninformative priors cause posterior impropriety, we assume a vague conjugate prior. Due to the complexity of the model, however, the posterior distributions are analytically intractable and hence a numerical scheme is required for the testing procedure. For this, we propose a Markov chain Monte Carlo (MCMC) algorithm to generate posterior samples of the parameters and use the samples to perform Bayesian testing of the bioequivalence. The proposed sampling-based approach also enables us to estimate the probability of any interesting properties of the parameters.

2. The Model and Hypotheses

For comparison of two drug formulations, a reference drug (R) and a test drug (T), consider the following model on a standard 2×2 crossover design,

$$y_{ijk} = \mu + G_k + S_{ik} + P_j + F_{(j,k)} + e_{ijk}, i = 1, \cdots, n_k;\ j = 1, 2;\ k = 1, 2, \quad (1)$$

where y_{ijk} is the response of the ith subject in the kth sequence for the jth period; μ is the overall mean; G_k is the fixed effect of the kth sequence with $\sum G_k = 0$; S_{ik} is the random subject effect of the ith subject in the kth sequence; P_j is the fixed effect of the jth period with $\sum P_j = 0$; $F_{(j,k)}$ is the direct fixed effect of the formulation in the kth sequence which is administered at the jth period such that $F_{(j,k)} = F_R$ for $(j,k) = (1,1), (2,2)$ and $F_{(j,k)} = F_T$ for $(j,k) = (1,2), (2,1)$ with $\sum\sum F_{(j,k)} = 0$; and e_{ijk} is the intra subject random error in observing y_{ijk}. It is assumed that S_{ik} are independent and identically distributed as $N(0, \sigma_s^2)$ and e_{ijk} are independently distributed as $e_{ijk} \sim N(0, \sigma_R^2)$ if $j = k$ and $e_{ijk} \sim N(0, \sigma_T^2)$ if $j \neq k$. Also, $\{S_{ik}\}$ and $\{e_{ijk}\}$ are assumed to be mutually independent. Since we are assumming normality, equivalence in distributions of the two drug bioavilabilities (PBE) can be established by demonstrating equivalence in both average and variability. Thus, PBE can be

assessed by testing the following interval hypotheses

$$H_0 : [\ |\ F_T - F_R\ | \leq \delta_1 \quad or \quad |\ F_T - F_R\ | \geq \delta_2] \quad or \quad [\frac{\sigma_T^2}{\sigma_R^2} \leq \epsilon_1 \ or \ \frac{\sigma_T^2}{\sigma_R^2} \geq \epsilon_2]$$
$$versus \qquad H_1 : \delta_1 \leq |\ F_T - F_R\ | \leq \delta_2 \ \ and \ \ \epsilon_1 \leq \frac{\sigma_T^2}{\sigma_R^2} \leq \epsilon_2, \tag{2}$$

where the constants δ_1, δ_2, ϵ_1 and ϵ_2 are chosen to define clinically significant differences. With these hypotheses, the bioequivalence in distribution between the two formulations is confirmed if H_0 is rejected. Thus, if H_0 has a smaller posterior probability than H_1 then H_0 would be rejected and the bioequivalence would be accepted.

3. Posterior Distribution

To compute the posterior probability of each hypothesis we need the joint posterior density function, $P(F_R, \sigma_T^2, \sigma_R^2 \mid \mathbf{y})$, of $(F_R, \sigma_T^2, \sigma_R^2)$, where $\mathbf{y} = \{y_{ijk}, i = 1, \cdots, n_k, j = 1, 2, k = 1, 2\}$. Parameters in model (1) are μ, $G_1 = -G_2$, $P_1 = -P_2$, $F_R = -F_T$, σ_T^2, σ_R^2 and σ_s^2. For simplicity, we denote $G = G_1$, $P = P_1$, $F = F_R$ and let $\rho = \sigma_s^2\sigma_R^2 + \sigma_s^2\sigma_T^2 + \sigma_R^2\sigma_T^2$ with constraint $\rho > \sigma_R^2\sigma_T^2$.

Note that (μ, G, P, F) are location parameters and $(\rho, \sigma_T^2, \sigma_R^2)$ are related to scale parameters. Assume prior independence between the location and scale parameters, the joint posterior density function of all the parameters can be given as

$$P(\mu, G, P, F, \rho, \sigma_T^2, \sigma_R^2 \mid \mathbf{y}) = P(\mu, G, P, F | \mathbf{y}, \rho, \sigma_T^2, \sigma_R^2) P(\rho, \sigma_T^2, \sigma_R^2 \mid \mathbf{y}). \tag{3}$$

The righthand side of (3) is a product of the joint conditional posterior density of the location parameters given the scale parameters and the joint posterior density function of the scale parameters.

Now, consider the joint conditional posterior density of the location parameters given the scale parameters. To derive the posterior distribution, we need to specify prior distribution of the parameters. For location parameters, uniform prior is most commonly used as a noninformative prior, hence we use the uniform prior $\pi(\mu, G, P, F) = 1$. Then, from the sufficient statistics for the location parameters and distributions of them (Grieve, 1985), it is easy to get

$$P(F \mid \mathbf{y}, \sigma_R^2, \sigma_T^2, \rho) \propto \frac{1}{\sigma_e} \exp[-\frac{8}{2m\sigma_e^2}(\hat{F} - F)^2], \tag{4}$$

where $\sigma_e^2 = (\sigma_R^2 + \sigma_T^2)/2$, $\hat{F} = \frac{1}{4}(\bar{y}_{.11} - \bar{y}_{.12} - \bar{y}_{.21} + \bar{y}_{.22})$ and $m = 1/n_1 + 1/n_2$.

Next, consider the joint posterior density function of the scale parameters. A 'two one-sided tests' procedure for assessment of bioequivalence in variability has been proposed by Liu(1991), based upon the idea of orthogonal transformation. The orthogonal transformation is as follows; Let $Y_{jk} = (Y_{1jk}, \cdots, Y_{n_k jk})^t$ and let c_{gk} be $n_k \times 1$ vector of coefficients of normalized linear orthogonal contrasts of degree n_k such that $1^t c_{gk} = 0$, $c_{gk}^t \cdot c_{gk} = 1$, $c_{gk}^t \cdot c_{g'k} = 0$ for $g \neq g'$, where $g = 1, \cdots, n_k - 1$; $j = 1, 2$, $k = 1, 2$. Let $Z_{gjk} = c_{gk}^t y_{jk} = c_{1gk}y_{1jk} + \cdots + c_{n_k gk}y_{n_k jk}$. Then,

$$Z = (Z_{111}, \cdot\cdot, Z_{(n_1-1)11}, Z_{122}, \cdot\cdot, Z_{(n_2-1)22}, Z_{121}, \cdot\cdot, Z_{(n_1-1)21}, Z_{112}, \cdot\cdot, Z_{(n_2-1)12})^t$$

is distributed as the multivariate normal distribution with mean 0 and covariance matrix Σ, where

$$\Sigma = \begin{pmatrix} (\sigma_s^2 + \sigma_R^2) I_{n_1+n_2-2} & \sigma_s^2 I_{n_1+n_2-2} \\ \sigma_s^2 I_{n_1+n_2-2} & (\sigma_s^2 + \sigma_T^2) I_{n_1+n_2-2} \end{pmatrix}.$$

Since the orthogonal transformation is linear, the units of the original data are maintained in the transformed data, and the likelihood function of Z is given as

$$\begin{aligned} f(z \mid \rho, \sigma_R^2, \sigma_T^2) \propto \quad & \rho^{-1/2} \exp[-\tfrac{1}{2\rho}[\tfrac{\rho+\sigma_T^4}{\sigma_R^2+\sigma_T^2}(\textstyle\sum z_{i11}^2 + \sum z_{i22}^2) \\ & + \tfrac{\rho+\sigma_R^4}{\sigma_R^2+\sigma_T^2}(\textstyle\sum z_{i21}^2 + \sum z_{i12}^2) \\ & - 2\tfrac{\rho-\sigma_R^2\sigma_T^2}{\sigma_R^2+\sigma_T^2}(\textstyle\sum z_{i11} z_{i21} + \sum z_{i22} z_{i12})]]. \end{aligned} \tag{5}$$

Now we need to specify the prior distribution of the scale parameters $(\rho, \sigma_T^2, \sigma_R^2)$. For the joint prior distribution of the scale parameters, either Jeffreys or reference prior is often used as a noninformative prior. In this case the reference prior would be more preferable since we have a dummy variable. The Jeffreys and the reference priors can be easily evaluated as

$$\pi(\sigma_R^2, \sigma_T^2, \rho) \propto \frac{1}{\rho^{3/2}(\sigma_R^2 + \sigma_T^2)} \quad \text{and} \quad \pi(\sigma_R^2, \sigma_T^2, \rho) \propto \frac{1}{\rho(\sigma_R^2 + \sigma_T^2)},$$

respectively. However, the induced posterior density functions from these priors are improper and hence the noninformative priors can not be used.

To get around this problem, we let $\pi(\rho) = 1/\rho$ and consider a proper prior for $\pi(\sigma_R^2, \sigma_T^2 \mid \rho)$ with restriction $\rho > \sigma_R^2\sigma_T^2$. For the prior $\pi(\sigma_R^2, \sigma_T^2 \mid \rho)$, noting that both the Jeffreys and the reference priors are functions of $\sigma_R^2 + \sigma_T^2$, we assume an inverse gamma distribution for $\sigma_e^2 = \frac{\sigma_R^2+\sigma_T^2}{2}$ and a uniform distribution on the interval $(0, 2\sigma_e^2)$ for σ_R^2 with the constraint $\rho > \sigma_R^2\sigma_T^2$. Note that σ_e^2 is the average of the variance corresponding to the reference drug and that to the test drug. With this form of prior, the posterior is proper.

From (4), (5) and the prior $\pi(\sigma_R^2, \sigma_T^2, \rho)$, one obtains

$$P(F, \sigma_R^2, \sigma_T^2, \rho \mid \mathbf{y}) \propto P(F \mid \sigma_R^2, \sigma_T^2, \rho, \mathbf{y}) P(\mathbf{z} \mid \sigma_R^2, \sigma_T^2, \rho) \pi(\sigma_R^2, \sigma_T^2, \rho). \tag{6}$$

Therefore, by integrating (6) with respect to ρ, the joint posterior density function of $(F, \sigma_R^2, \sigma_T^2)$ is obtained.

However, it is impossible to compute the posterior probabilities of hypotheses from the posterior density function because it is highly complicated. Thus, we need to rely on numerical schemes to estimate the posterior probabilities. We propose to use a Markov chain Monte Carlo (MCMC) algorithm to generate posterior samples of the parameters and estimate the posterior probabilities by using the samples. MCMC methods have been successful for solving complicated problems in wide applications of statistics.

4. A MCMC Algorithm and Bayes Test

If we let $u = \sum z_{i11}^2 + \sum z_{i22}^2$, $v = \sum z_{i21}^2 + \sum z_{i12}^2$, $w = \sum z_{i11}z_{i21} + \sum z_{i22}z_{i12}$ for notational convenience and let $\sigma_e^2 = (\sigma_R^2+\sigma_T^2)/2$, the joint posterior of $(F, \sigma_R^2, \sigma_e^2)$ becomes, by change of variables,

$$P(F, \sigma_R^2, \sigma_e^2 \mid \mathbf{y}) \propto \frac{1}{\sqrt{\sigma_e^2}} \exp[-\frac{4}{m\sigma_e^2}(\hat{F} - F)^2] \tag{7}$$

$$\times \int_{\sigma_R^2(2\sigma_e^2-\sigma_R^2)}^{\infty} \rho^{-3/2} \exp[-\frac{1}{4\rho\sigma_e^2}(u(2\sigma_e^2 - \sigma_R^2)^2 + v\sigma_R^4 + 2w\sigma_R^2(2\sigma_e^2 - \sigma_R^2)^2)]d\rho \tag{8}$$

$$\times(\frac{1}{2\sigma_e^2})^{\alpha+1} \exp[-\frac{1}{2\sigma_e^2}(\frac{u+v-2w}{2} + \frac{2}{\beta})]\frac{1}{2\sigma_e^2} I_{(2\sigma_e^2 > \sigma_R^2)}. \tag{9}$$

Note that (7) is the conditional posterior density function of F given (σ_R^2, σ_e^2) and (8) and (9) are the posterior density function of (σ_R^2, σ_e^2).

The conditional posterior distribution of F is $N(\hat{F}, m\sigma_e^2/8)$ and hence sample generation from this distribution is straightforward. We suggest to generate samples of (σ_R^2, σ_e^2) by using a Metropolis-Hastings (M-H) algorithm and then generate F from its conditional posterior distribution. For the proposal density in the M-H algorithm, we propose to use $g_1(\sigma_e^2)g_2(\sigma_R^2|\sigma_e^2)$, where $g_1(\sigma_e^2)$ is the density function of $Inverse\ Gamma(\alpha, (\frac{u+v-2w}{4} + \frac{1}{\beta})^{-1})$ and $g_2(\sigma_R^2|\sigma_e^2)$ is the density function of $Uniform(0, 2\sigma_e^2)$. Note that $g_1(\sigma_e^2)g_2(\sigma_R^2|\sigma_e^2)$ is equivalent to (9) and hence the weight function in the M-H algorithm is given as (8). From the MCMC iterations, samples from a burn-in phase of the Markov chain is discarded and samples from iterations after the burn-in can be taken as posterior samples of the parameters. From these samples, we can estimate posterior probabilities of H_0 and H_1 by the relative frequencies of samples satisfying H_0 and H_1, respectively. With no prior preference on either hypothesis, we can conclude that the two formulations are bioequivalent in terms of population bioequivalence, if the estimated posterior probability of H_1 is larger than 0.5.

5. An Example

To illustrate application of the proposed statistical test for assessment of population bioequivalence, we consider an example concerning bioavailability between two formulations which were considered by Liu and Chow(1992). The experiment in the example was conducted by a standard 2×2 crossover design and it was shown that the two drug formulations in the example are bioequivalent in average but not in variability (Chow and Liu,1992; Liu and Chow,1992; Chen et al.,1996). From the data, we get $m = 1/6$, $\hat{F} = 1.14375$ and MSE estimator of σ_e^2 is $\hat{\sigma_e^2} = 167.25$ under the assumption $\sigma_R^2 = \sigma_T^2$. Also, we get $u = 9692.70$, $v = 10198.3$ and $w = 6266.03$ after the orthogonal transformation. Hence a 'best guess' for σ_e^2 would be 167.25 and, from the 95% confidence interval for σ_e^2 which is $(100.037, 335.032)$, the upper

$97.5th$ percentile can be chosen as 335.032. Thus, following the guidlines given in Section 3, we choose $\hat{\alpha} = 7$ and $\hat{\beta} = 0.001$. The bioequivalence limits are chosen according to ± 20 rule for illustrative purposes.

From 10,000 random samples generated from the algorithm, we obtain 10.57% ($< 50\%$) as the estimated posterior probability of H_1. Hence H_0 is not rejected and we fail to assess population bioequivalence for the two formulations. Seperate tests of bioequivalence in mean and variability yield 98.84% and 10.65% as the estimated posterior probabilities of H_1, respectively, showing that the two formulations are bioequivalent in mean but not in variability. This coincides with the results of the classical tests by Chow and Liu(1992), Liu and Chow(1992), and Chen et al.(1996).

Now, if we take $\epsilon_1 = 0$, that is, if we treat any amount of reduction in variability of the Test drug relative to the Reference drug as being consistent with population bioequivalence (Hauck et al., 1997), then we would get a bigger estimate of the posterior probability of H_1. In this example, we get 53.67% as the estimated posterior probability of H_1 from simulation (with $\epsilon_2 = 1.2$), hence H_0 is rejected. If we use ϵ_2 bigger than 1.2, for instance 1.5 as in Hauck et al.(1997), the result would be more convincing for PBE of the two formulations. In our example, we get 59.21% as the estimated posterior probability of H_1 with $\epsilon_2 = 1.5$.

REFERENCES

Anderson, S. and Hauck, W.W. (1990). Consideration of individual bioequivalence. *Journal of Pharm. & Biopharm.*, 18, 259-273.

Chen, K-W., Li, G., Sun, Y., and Chow, S-C. (1996). A confidence region approach for assessing equivalence in variability of bioavailability. *Biometrical Journal*, 38, 475-487.

Chow, S-C. (1999). Individual bioequivalence - a review of the FDA draft guidance. *Drug Information Journal*, 33, 435-444.

Chow, S-C. and Liu, J-P. (1992). Design and analysis of bioavailability and bioequivalence, New York, NY: Marcel Dekker, Inc.

Grieve, A. P. (1985). A Bayesian analysis of the two-period crossover design for clinical trials, *Biometrics*, 41 979-990.

Hastings, W.K. (1970). Monte Carlo sampling methods using markov chains and their applications, *Biometrika*, 57 97-109.

Hauck, W.W., Bois, F.Y., Hyslop, T., Gee, L. and Anderson, S. (1997). A parametric approach to population bioequivalence, *Statistics in Medicine*, 16, 441-454.

Liu, J-P. (1991). Bioequivalence and intrasubject variability, *Journal of Biopharmaceutical Statistics*, 1 205-219.

Liu, J-P. and Chow, S-C. (1992). On the assessment of variability in bioavailability/bioequivalence studies, *Communications in Statistics-Theory and Methods*, 21 2591-2607.

Schall, R. (1995). Assessment of individual and population bioequivalence using the probability that bioavailabilities are similar. *Biometrics*, 51, 615-626.