STATISTICAL TESTS FOR MULTIVARIATE BIOEQUIVALENCE

by

Anirban DasGupta
Purdue University

Weizhen Wang
Wright State University

J.T. Gene Hwang
Cornell University

Technical Report #96-39

Department of Statistics
Purdue University

September 1996

*Research supported by NSF Grant DMS 9307727

STATISTICAL TESTS FOR MULTIVARIATE BIOEQUIVALENCE

by

Weizhen Wang
Dept. of Mathematics and Statistics
Wright State University
Dayton, OH 45435

and

Anirban DasGupta
Dept. of Statistics
Purdue University
West Lafayette, IN 47907

J.T. Gene Hwang
Department of Mathematics
Cornell University
Ithaca, NY 14853

Abstract

Although the FDA recommends testing bioequivalence of individual pharmacokinetic parameters one at a time, it seems more reasonable to conduct a test simultaneously for all the parameters. In this paper, we discuss several ways to contruct such tests. It is shown that the confidence set approach leads to a test which can be uniformly improved by the intersection of Schuirmann's two one-sided tests procedure. The latter test can further be improved upon substantially by using the one-dimensional unbiased test of Brown, Hwang and Munk (1996). Numerical calculations of powers are given to support this claim.

Statistical Tests for Multivariate Bioequivalence

Weizhen Wang
Department of Mathematics and Statistics
Wright State University
Dayton, OH 45435

Anirban DasGupta
Department of Statistics
Purdue University
West Lafayette, IN 47907

J. T. Gene Hwang
Department of Mathematics
Cornell University
Ithaca, NY 14853

September 9, 1996

Abstract

Although the FDA recommends testing bioequivalence of individual pharmacokinetic parameters one at a time, it seems more reasonable to conduct a test simultaneously for all the parameters. In this paper, we discuss several ways to construct such tests. It is shown that the confidence set approach leads to a test which can be uniformly improved by the intersection of Schuirmann's two one-sided tests procedure. The latter test can further be improved upon substantially by using the one-dimensional unbiased test of Brown, Hwang and Munk (1996). Numerical calculations of powers are given to support this claim.

KEY WORDS: Likelihood ratio test; Intersection-Union method; Confidence set.

1 INTRODUCTION

In bioequivalence studies, one typically is interested in demonstrating that a new drug is similar in efficacy to a brand-named drug. The FDA (1992) recommends a 2×2 crossover design. Typically twenty four subjects are randomly divided into two groups. One group will be applied the brand-named drug and after a washout period the new drug. The other group is similarly treated except that the order of drugs is reversed. The blood samples are then collected from each subject at various times and a blood concentration curve against time of a certain ingredient is obtained.

The three most typical characteristics of the blood concentration curve considered are the area under the concentration curve (AUC), the maximum concentration (C_{max}) and the time to reach the maximum concentration (T_{max}). FDA (1992) then recommends that one applies Schuirmann's (1987) two one-sided tests procedure to test the hypotheses:

$$H_0 : |\theta| \geq \Delta \ \text{ vs } \ H_A : |\theta| < \Delta.$$

Here $\theta = \theta_T - \theta_R = ln(\mu_T/\mu_R)$, where ln is the natural logarithm function, and θ_T and θ_R represent respectively the means of the characteristic in a ln scale corresponding to the new treatment and the reference (branded name) treatment. The FDA's recommended cutoff number Δ is $ln(1.25)$ so that μ_T and μ_R stay within 80% of each other. If H_0 is rejected, then bioequivalence is declared.

Schuirmann's test, however, is one-dimensional. It would seem more reasonable to consider simultaneously the test involving all relevant characteristics. In this paper we shall consider testing the hypotheses:

$$H_0 : \max_{1 \leq i \leq p} |\underline{\theta}^{(i)}| \geq \Delta \ \text{ vs } \ H_A : \max_{1 \leq i \leq p} |\underline{\theta}^{(i)}| < \Delta, \qquad (1.1)$$

instead. Here $\underline{\theta}^{(i)}$ relates to, for example, AUC, C_{max} and T_{max}. Under a fairly general linear model including the crossover design as a special case, we may consider the canonical form

$$X_j \overset{i.i.d}{\sim} N_p(\underset{\sim}{\theta}, \Sigma), j = 1, ..., n, \tag{1.2}$$

where X_j represents an unbiased estimate of $\underset{\sim}{\theta} = (\theta^{(1)}, ..., \theta^{(p)})'$, n is the number of subjects, p is the number of characteristics, and Σ is the $p \times p$ unknown covariance matrix. The sufficient statistics are the sample mean and the sample covariance matrix:

$$\bar{X} = \frac{\sum_{j=1}^{n} X_j}{n}, \quad \hat{\Sigma} = \frac{\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})'}{n-1}, \tag{1.3}$$

where, for any vector V, V' denotes the transpose of V. This paper focuses on deriving tests for (1.1).

When $p = 1$, i.e., only one characteristic is used in analysis, Schuirmann (1987) provided an $\alpha-$level test for (1.1) by using the intersection-union method with the rejection region:

$$|\bar{X}| < \Delta - t_{n-1}(\alpha)\hat{\Sigma}^{1/2}/\sqrt{n}, \tag{1.4}$$

where $t_{n-1}(\alpha)$ is the upper α quantile of Student t distribution with $n-1$ degrees of freedom. Brown et al. (1996) obtained an $\alpha-$level unbiased test with the rejection region:

$$|\bar{X}| < B(\hat{\Sigma}^{1/2}), \tag{1.5}$$

where B is some positive function which is always no smaller than the right hand side of (1.4). Therefore, Brown et al.'s test (1.5) uniformly improves Schuirmann's test (1.4) in power. This fact will be used in Section 5. Although there is no closed form for the function B, it can be evaluated numerically for (1.1).

In Section 2, we shall use the likelihood ratio approach to derive a test statistic. Although the likelihood ratio test (LRT) generally satisfies a well known asymptotic result to help choose the cutoff point, it fails to recommend a meaningful cutoff point for (1.1). In Section 3, we shall show that a confidence set approach leads to the same statistic and a specific cutoff point. Section 4 shows that the intersection-union method (see, for example, Berger (1982)) leads to yet another test, called the intersection-union test, which uses the same statistic and a smaller cutoff point when compared to the test derived by the confidence set approach. Consequently, the α-level intersection-union test is uniformly more powerful than the test constructed using the (1-α) confidence set. Similar phenomenon has been observed in the one-dimensional case. It is well known that the usual $(1-\alpha)$ t-confidence interval leads to a test for (1.1) which is only of size of $\alpha/2$ and is uniformly less powerful than the two one-sided tests procedure of size α. Here and later, the size of a test is defined to be the supremum of the type I error over the null hypothesis space. In contrast, a test has level α if its size is no greater than α. The intersection-union test, however, can be improved uniformly as shown in section 5. By adapting the test of Brown et al. (1996), we can construct a multivariate test uniformly improving upon the intersection-union test. The improvement in power may be as big as 0.11. Numerical studies are reported in Section 6 before the conclusion in Section 7.

2 LIKELIHOOD RATIO TEST

The likelihood ratio approach is one of the most common ways of constructing a test. Let $L(\underline{X}, \underline{\theta}, \Sigma)$ be the likelihood function of model (1.2), where $\underline{X} = (X_1, ..., X_n)$. The test statistic is defined to be

$$\lambda(\underline{X}) = \frac{\sup_{H_0} L(\underline{X}, \underline{\theta}, \Sigma)}{\sup_{H_0 \cup H_A} L(\underline{X}, \underline{\theta}, \Sigma)} \tag{2.1}$$

and we reject the null hypothesis if $\lambda(\underline{X}) < K$ for some suitable K. It is interesting that (2.1) has a simple expression which leads to the simple test in Theorem 1 below. In comparison the likelihood ratio test for testing H_A against H_0 has a very complicated form especially for large p and when Σ is nondiagonal since it involves calculating the maximum likelihood estimator for θ when $\theta \in H_A$, a calculation known to have only numerical solutions. See the comment at the end of the proof of Theorem 1 in the Appendix.

Lemma 1 *Under the model (1.2),*

$$\lambda(\underline{X}) = \left(\frac{1}{1 + inf_{\underline{\theta} \in H_0} n(\bar{X} - \underline{\theta})' \hat{\Sigma}^{-1} (\bar{X} - \underline{\theta})/(n-1)}\right)^{n/2}. \tag{2.2}$$

Proof. See the appendix.

Below let $\bar{X}^{(i)}$ denote the ith element of $\bar{X}$ and $\hat{\Sigma}_{ii}$ the ith diagonal element of $\hat{\Sigma}$.

Theorem 1 *Consider the testing problem (1.1) under model (1.2). Then, for* $0 < K < 1$,

$$\lambda(\underline{X}) < K \quad \textit{if and only if } |\bar{X}^{(i)}| < \Delta - C\sqrt{\hat{\Sigma}_{ii}/n} \ \ \forall \ 1 \le i \le p \tag{2.3}$$

where $C^2 = K^{-2/n} - 1$.

Proof. See the appendix.

There is one problem in this approach, i.e., the determination of the critical value K. Traditionally, one may conclude that $-2ln(\lambda(\underline{X}))$ has an asymptotic chi-squared distribution with degrees of freedom $d - d_{H_0}$, where d is the dimension of the whole parameter space and d_{H_0} is the dimension of the null hypothesis space. In testing (1.1), however, $d - d_{H_0} = 0$ which is nonsensical since a chi-squared distribution with zero degree does not exist. However, the likelihood ratio test does give us a useful statistic. We postpone to Section 4 the determination of the critical value C so that the likelihood ratio test is an α-level test.

3 CONFIDENCE SET APPROACH

Another way of constructing a test is the confidence set approach. Earlier Hsu, Lu and Chan (1995) proposed using two separate confidence sets, one for the new and one for the reference treatment. Their approach should be less efficient than considering one confidence set for the difference as we shall do below. Brown, Casella and Hwang (1995) constructed confidence sets which are optimal in some sense for bioequivalence problems. They assume that Σ is known, although it is possible to generalize their result to the unknown Σ case.

For testing a statistical hypothesis, let $C(\underline{X})$ be a confidence set for $\underline{\theta}$. One may define a test for (1.1) which rejects the null hypothesis if and only if

$$C(\underline{X}) \subset H_A. \tag{3.1}$$

It is easy to see that this test would be an α-level test if $C(\underline{X})$ has confidence level $1-\alpha$. We shall consider the Hotelling's $T^2 = n(\bar{X}-\underline{\theta})'\hat{\Sigma}^{-1}(\bar{X}-\underline{\theta})$. Then $(n-p)T^2/[(n-1)p]$ has an $F-$distribution with degrees of freedom p and $n-p$. Let $F_{p,n-p}(\alpha)$ be its upper α quantile and C_1 be the value such that $P(T^2 \leq C_1^2) = 1-\alpha$, and

$$C(\underline{X}) = \{\underline{\theta} : T^2 \leq C_1^2\}. \tag{3.2}$$

Then $C(\underline{X})$ is a $1-\alpha$ confidence set for $\underline{\theta}$. In fact,

$$C_1^2 = F_{p,n-p}(\alpha)\frac{p}{n-p}(n-1). \tag{3.3}$$

Theorem 2 *Consider the testing problem (1.1) under model (1.2). Then the test defined by (3.1) and (3.2) is equal to the test which rejects the null hypothesis of (1.1) if and only if*

$$|\bar{X}^{(i)}| < \Delta - C_1\sqrt{\hat{\Sigma}_{ii}/n} \ \ \forall\ 1 \leq i \leq p \tag{3.4}$$

Proof. See appendix.

Notice that this test uses the same statistic as the likelihood ratio test (2.3). It also gives a precise cutoff point as shown. This is a very useful feature of this test. It can, however, be uniformly improved by the intersection-union test in the next section.

4 INTERSECTION-UNION TEST

Schuirmann (1987) proposed a test (1.4) by using an approach called intersection-union method in Berger (1982). See also Casella and Berger (1990, p.356). In this section we will generalize the Schuirmann's test to $p > 1$.

Let us consider p sets of hypotheses:

$$H_{0i} : |\underline{\theta}^{(i)}| \geq \Delta \quad H_{Ai} : |\underline{\theta}^{(i)}| < \Delta, \tag{4.1}$$

for $i = 1, ..., p$. For each set of hypotheses, one has the α-level rejection region

$$R_i = \{|\bar{X}^{(i)}| < \Delta - t_{n-1}(\alpha)\sqrt{\hat{\Sigma}_{ii}/n}\}$$

corresponding to the Schuirmann's test (1.4) for the one-dimensional case. It is obvious that the rejection region

$$R^I = \cap_{i=1}^{p} R_i \tag{4.2}$$

defines an α-level test for (1.1) since $H_A = \cap_{i=1}^{p} H_{A_i}$ and R^I has a form similar to the likelihood ratio test (2.3). This method is called intersection-union since the rejection region is the intersection of several rejection regions and the null hypothesis is the union of several null hypotheses. This test is also called the repeated univariate test (based on decision rule 2) in Schall et al. (1996).

Theorem 3 *Consider the testing problem (1.1) under model (1.2). Then the test R^I for (1.1) has size α. It is uniformly more powerful than the test (3.4).*

Table 1: The actual size α_1 of the test derived from the confidence set approach when the test level is 0.05, the sample size $n = 24$ and the number of characteristics p varies.

p	1	2	3	4	5	10
α_1	0.025	6.63×10^{-3}	2.10×10^{-3}	6.70×10^{-4}	2.35×10^{-4}	3.68×10^{-7}

Proof. See appendix.

From Theorem 3, we can determine the critical value $C = t_{n-1}(\alpha)$ in (2.3) so that the likelihood ratio test (2.3) has size α. Therefore the α-level likelihood ratio test is equal to the intersection-union test. Also, we can figure out the actual size of the test (3.4) constructed by the confidence set approach. The size is the number α_1 such that $C_1 = t_{n-1}(\alpha_1)$. After some straightforward calculations, we may conclude that

$$\alpha_1 = P(\frac{\chi_1^2}{\chi_{n-1}^2} > q)/2 \tag{4.3}$$

where q is the α upper quantile of χ_p^2/χ_{n-p}^2. Here χ_p^2 denotes a chi-squared random variable with p degrees of freedom, and all the above chi-squared random variables are independent.

In particular, if $p = 1$, then $\alpha_1 = \alpha/2$, a well known result in the one-dimensional case. Namely the $1 - \alpha$ confidence interval leads to the two one-sided tests procedure with size $\alpha/2$. In general, we can conclude that

$$\alpha_1 \leq \alpha/2$$

for all p. Also α_1 is decreasing and can be as close to zero as possible when p goes to infinity. Table 1 gives the value of α_1 and as is shown, α_1 can be very small. This shows that the discrepancy between the error probability α of the confidence set can be drastically different from the actual size of the test that the confidence set leads to. This striking phenomenon is unique to the bioequivalence problem, since in standard problems the error probability of a confidence interval equals the size of its associated test.

5 IMPROVED TEST

Although the intersection-union test R^I improves upon the test constructed using the confidence approach, it can be uniformly improved in power itself. In this section we will generalize Brown et al.'s (1996) test (1.5) to the case $p > 1$. Here, we will apply the intersection-union method again.

Consider the rejection region of an unbiased test

$$R_i^U = \{|\bar{X}^{(i)}| < B(\sqrt{\hat{\Sigma}_{ii}})\}$$

of Brown et al. (1996) (1.5) for the one-dimensional case. For the specific form of B, see Brown et al. (1996). It is known that R_i^U is an unbiased α-level test for (4.1) which contains properly the rejection region of Schuirmann's test. Consequently, it has a uniformly larger power than Schuirmann's test. Let

$$R^U = \cap_{i=1}^p R_i^U.$$

Obviously R^U uniformly improves R^I. Further, by an argument similar to the one in the last section, we conclude that R^U has size α for testing (1.1). The test R^U is not unbiased for (1.1) unless $p = 1$. It might be uniformly improved in power by enlarging the rejection region without overshooting the test level. The problem of constructing a test better than R^U is still unsolved. Both R^U and R^I have the good property that their power functions are decreasing in each $|\underline{\theta}^{(i)}|$ when other $\underline{\theta}^{(j)}$'s and Σ are fixed. This is reasonable since we want to have a maximum power when $\underline{\theta} = 0$, i.e., the two drug effects have the same mean. Unlike the one-dimensional case, when $p > 1$ we shall show below that a nontrivial unbiased test with such a power function does not exist. Therefore, an unbiased test for (1.1), if exists, may have a pathological rejection region and may be difficult to construct.

Theorem 4 *Let $\Phi(\bar{X}, \hat{\Sigma})$ be the critical function of an α-level unbiased test for (1.1) whose power function is unimodal with respect to each coordinate $\underline{\theta}^{(i)}$. Then $\Phi(\bar{X}, \hat{\Sigma})$ is the trivial test, i.e., $\Phi(\bar{X}, \hat{\Sigma}) = \alpha, a.s.$.*

Proof. See appendix.

6 NUMERICAL STUDIES OF POWER

In this section we will provide some simulation results for the power of the derived tests in Sections 4 and 5. We choose $p = 3$ and $\Delta = ln(1.25)$ in (1.1) which means that we may consider AUC, C_{max} and T_{max} simultaneously and the ratios of two population means for the characteristics must be within $(0.8, 1.25)$ for bioequivalence to be asserted.

Now the power is a function of

$$\underline{\theta} = \begin{pmatrix} \underline{\theta}^{(1)} \\ \underline{\theta}^{(2)} \\ \underline{\theta}^{(3)} \end{pmatrix}, \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}$$

in (1.2), and σ_i is the standard deviation of each variable and ρ_{ij} is the correlation coefficient of each pair of variables. In the simulation, $\underline{\theta}^{(i)} = a, \sigma_i = b$ and $\rho_{ij} = c$ for all i, j. It is reasonable to assume that ρ_{ij}'s are positive since AUC, C_{max} and T_{max} usually are positively associated. All results are based on 100,000 simulation times.

Table 2 gives the power simulations of the intersection-union test and the improved test in Section 5 when $a = 0, b = 0.5$ and c is from 0 to 1 with a step 0.1. The improved test, as expected, has a uniformly higher power than the intersection-union test (4.2). The power increment is getting larger and larger when the correlation coefficient goes to 1. When $b = 0.6$ and $c = 1$, the improved test can have a power 0.244 which

Table 2: The power simulations of the intersection-union test (IUT) and the improved test in Section 5 when $a = 0, b = 0.5$ or 0.6 and c varies.

	b=0.5		b=0.6	
c	IUT	Improved test	IUT	Improved test
0	0.04893	0.06557	0.00291	0.01446
0.1	0.05183	0.06875	0.002470	0.01516
0.2	0.05478	0.06991	0.002880	0.01497
0.3	0.05825	0.07539	0.003630	0.01601
0.4	0.06379	0.08278	0.004730	0.01870
0.5	0.07440	0.09231	0.006170	0.02136
0.6	0.08849	0.10814	0.008610	0.02502
0.7	0.10908	0.12864	0.011920	0.03169
0.8	0.13960	0.16192	0.020110	0.04516
0.9	0.19367	0.21815	0.037560	0.07271
1	0.37123	0.40705	0.13811	0.24431

is 0.11 larger than the intersection-union test. Also one may notice that the power increment is sensitive to the correlation coefficient. For example, when $b = 0.5$ and c is changed from 0.9 to 1, the power of the improved test increases from 0.218 to 0.407. Therefore, it is much easier to establish the bioequivalence if the characteristics are extremely highly correlated.

In Table 2, we focus on the case where powers are not high. These are the cases where the improvements are largest. When powers are high, the improvement is negligible. However, since the improved test is uniformly more powerful, there is no loss in any event but possible with some gain in using the improved test.

7 CONCLUSION

In this paper we discuss several different ways of constructing bioequivalence test in a multivariate setting when the covariance matrix is unknown. It is interesting to observe that the likelihood ratio approach, the confidence set approach and the intersection-union (IU) approach all lead to the same statistic. However, the first does not provide a cutoff point; and the IU approach gives a better cutoff point than the confidence set approach and hence provides a more powerful test. This IU approach can be uniformly improved by using Brown et al's (1996) one-dimensional result. It appears that it is difficult to further improve upon this improved test. In particular, under a certain condition, any unbiased test for $p > 2$ is a trivial test (see Theorem 4).

8 APPENDIX

Proof of Lemma 1.

$$L(\underline{\theta}, \Sigma|\underline{X}) = \frac{1}{(\sqrt{2\pi})^{np}|\Sigma|^{n/2}}\exp\{-\sum_{j=1}^{n}(X_j - \underline{\theta})'\Sigma^{-1}(X_j - \underline{\theta})/2\}.$$

Here, for a matrix M, $|M|$ denotes its determinant. It is well known that

$$\sup{}_{H_0 \cup H_A} L(\underline{\theta}, \Sigma|\underline{X}) = L(\underline{\theta}, \Sigma|\underline{X})|_{\underline{\theta}=\bar{X}, \Sigma=(n-1)\hat{\Sigma}/n} = \frac{e^{-n/2}}{(\sqrt{2\pi})^{np}|(n-1)\hat{\Sigma}/n|^{n/2}}.$$

To evaluate the numerator in (2.1), note that

$$\sup{}_{\Sigma} L(\underline{\theta}, \Sigma|\underline{X}) = \frac{e^{-n/2}}{(\sqrt{2\pi})^{np}|\sum_{j=1}^{n}(X_j - \underline{\theta})(X_j - \underline{\theta})'/n|^{n/2}}, \tag{8.1}$$

and

$$\begin{aligned} |\sum_{j=1}^{n}(X_j - \underline{\theta})(X_j - \underline{\theta})'| &= |(n-1)\hat{\Sigma} + n(\bar{X} - \underline{\theta})(\bar{X} - \underline{\theta})'| \\ &= |(n-1)\hat{\Sigma}||(I + n\hat{\Sigma}^{-1/2}(\bar{X} - \underline{\theta})(\bar{X} - \underline{\theta})'\hat{\Sigma}^{-1/2}/(n-1)| \\ &= |(n-1)\hat{\Sigma}|(1 + n(\bar{X} - \underline{\theta})'\hat{\Sigma}^{-1}(\bar{X} - \underline{\theta})/(n-1)). \end{aligned}$$

Hence (8.1) equals

$$\frac{e^{-n/2}}{(\sqrt{2\pi})^{np}|(n-1)\hat{\Sigma}/n|^{n/2}(1 + n(\bar{X} - \underline{\theta})'\hat{\Sigma}^{-1}(\bar{X} - \underline{\theta})/(n-1))^{n/2}},$$

establishing (2.2).

Proof of Theorem 1. By Lemma 1, $\lambda(\underline{X}) \leq K$ is equivalent to

$$\inf_{\underline{\theta}\in H_0} d(\bar{X},\underline{\theta}) > C^2. \tag{8.2}$$

where $d(\bar{X},\underline{\theta}) = (\bar{X}-\underline{\theta})'\hat{\Sigma}^{-1}(\bar{X}-\underline{\theta})$.

If $\bar{X} \in H_0$, then (8.2) fails. Since by setting $\bar{X} = \underline{\theta}$, we may see that the left hand side of (8.2) is zero. Obviously (2.3) fails also. Hence (2.3) and (8.2) are equivalent for this trivial case.

If $\bar{X} \notin H_0$, then $\bar{X} \in H_A$ and hence $\bar{X}$ stays inside a cube. We shall obtain an expression for the left hand side of (8.2), by finding a $\underline{\theta}_{min}$ outside the cube such that $d(\bar{X},\underline{\theta})$ is minimized at $\underline{\theta} = \underline{\theta}_{min}$. Let P_{ij} denote the $(p-1)$-dimensional plane which consists of points whose ith coordinate is equal to $j\Delta$ where $j = 1$ or $j = -1$. Obviously these P_{ij}, $(1 \leq i \leq p$ and $j = \pm 1)$, contain the boundary of the cube. Now let

$$d_{ij} = min_{\underline{\theta}\in P_{ij}} d(\bar{X},\underline{\theta}).$$

We claim that

$$d^* = min_{i,j} d_{ij}. \tag{8.3}$$

where $d^* = d(\bar{X},\underline{\theta}_{min})$ is the global minimum within $\underline{\theta} \in H_0$, i.e., d^* equals the left hand side of (8.2). Equation (8.3) will be proved after (8.6). Let us assume it for the time being. Note that P_{ij} consists of point $\underline{\theta}$ such that

$$\underline{\theta}' e_i = j\Delta$$

where e_i is the ith coordinate vector. Let $\eta = \hat{\Sigma}^{-1/2}\underline{\theta}$. Using the new notation, we may write P_{ij} as

$$\eta'\hat{\Sigma}^{1/2} e_i = j\Delta \tag{8.4}$$

and the distance as

$$d(\bar{X},\underline{\theta}) = (\hat{\Sigma}^{-1/2}\bar{X} - \eta)'(\hat{\Sigma}^{-1/2}\bar{X} - \eta) \tag{8.5}$$

The minimization problem is now equivalent to minimizing (8.5) over all η satisfying (8.4). This is equivalent to find the minimum distance between $\hat{\Sigma}^{-1/2}\bar{X}$ and the plane defined by (8.4). Note that the distance between a point y_0 and the equation $y'a = b$ is

$$\frac{a'y_0 - b}{\sqrt{a'a}}.$$

Letting $a = \hat{\Sigma}^{1/2}e_i$ and $b = j\Delta$, then

$$d_{ij}^2 = \frac{|\bar{X}^{(i)} - j\Delta|^2}{\hat{\Sigma}_{ii}}.$$

Using this and (8.4), (8.2) is then equivalent to

$$\frac{|\bar{X}^{(i)} - j\Delta|}{\hat{\Sigma}_{ii}^{1/2}} < C. \tag{8.6}$$

Combining the two inequalities for $j = \pm 1$, (8.6) is equivalent to (2.3), establishing Theorem 1.

To prove (8.3), we first note that obviously

$$d^* \leq d_{ij}$$

for every i, j, since P_{ij} is a subset of H_0. Hence

$$d^* \leq min_{i,j} d_{ij}.$$

Further, the global minimum d^* is achieved on the boundary of the cube. This is due to the fact that the minimizer $\underline{\theta}_{min}$ is the point of the intersection with the cube of the largest ellipsoid of the form

$$S = \{\underline{\theta} : d(\bar{X}, \underline{\theta}) = \text{constant}\}$$

such that S is contained in the cube. Using this, we then have

$$d^* \geq min_{i,j} d_{ij}$$

since the boundary of the cube is contained in the union of Pij's. Hence (8.3) is established and the proof is completed.

A Comment on Theorem 1: Note that we are fortunate to have a simple expression for d^* by using (8.3). There is no simple expression for

$$\min_{\underline{\theta} \in H_A} d(\bar{X}, \underline{\theta}),$$

however, when $\bar{X}$ is outside the cube H_A. This kind of minimization relates to the problem of finding the maximum likelihood estimator of $\underline{\theta}$ when $\underline{\theta}$ is known to be in a cube, which generally cannot be found in closed form. As a result, only numerical solutions are possible.

Proof of Theorem 2. Let $b(\underline{X}) = \max_{1 \le i \le p} b_i(\underline{X})$, where $b_i(\underline{X}) = \sup_{\underline{\theta} \in C(\underline{X})} |\underline{\theta}^{(i)}|$. By (3.3), the test rejects if $b(\underline{X}) < \Delta$. We claim that

$$b_i(\underline{X}) = |\bar{X}^{(i)}| + C_1\sqrt{\hat{\Sigma}_{ii}/n} \tag{8.7}$$

which obviously establishes the theorem.

To prove (8.7), let e_i be as in the proof of Theorem 1. Then $\underline{\theta}^{(i)} = e_i'\underline{\theta}$, and for $\underline{\theta} \in C(\underline{X})$

$$\begin{aligned} e_i'\underline{\theta} &\le \bar{X}^{(i)} + (\hat{\Sigma}^{1/2}e_i)'\hat{\Sigma}^{-1/2}(\underline{\theta} - \bar{X}) \\ &= \bar{X}^{(i)} + \sqrt{e_i'\hat{\Sigma}e_i[(\underline{\theta} - \bar{X})'\hat{\Sigma}^{-1}(\underline{\theta} - \bar{X})]} \\ &\le \bar{X}^{(i)} + C_1\sqrt{e_i'\hat{\Sigma}e_i}/\sqrt{n} \\ &= \bar{X}^{(i)} + C_1\sqrt{\hat{\Sigma}_{ii}/n} \end{aligned}$$

where the first inequality follows from the Cauchy-Schwartz inequality. Similarly, $e_i'\underline{\theta} \ge \bar{X}^{(i)} - C_1\sqrt{\hat{\Sigma}_{ii}/n}$. These imply that $b_i(\underline{X})$ is less than or equal to the right hand side of (8.7). To establish the equality, let

$$\underline{\theta} = \bar{X} \pm C_1\hat{\Sigma}e_i/\sqrt{\hat{\Sigma}_{ii}n}.$$

It can be shown that these two $\underline{\theta}'s$ belongs to $C(\underline{X})$ and also one of them gives equality in (8.7). This establishes the theorem.

Proof of Theorem 3. By the intersection-union method, R^I is an α-level test. Now we show the supremum of the probability of type I error over H_0 is at least α. These obviously imply that the size of the test is exactly α. To achieve this, let $\underline{\theta}^{(i)}$ be a fixed constant for every i and let all entries in Σ be the same. Therefore, all coordinates of X_i in model (1.2) are the same, and thus R^I reduces to the Schuirmann's test (1.4). It is well known that the Schuirmann's test has size α, establishing the theorem.

To prove the second part of the theorem, we show that $t_{n-1}(\alpha)$ is less than C_1 which will imply that R^I is uniformly more powerful than (3.4). This follows from (3.3) and the fact, which is easily proved, that the random variable $p(n-1)F_{p,n-p}/(n-p)$ is stochastically greater than the random variable $F_{1,n-1}$, where $F_{p,q}$ has an F-distribution with degrees of freedom p and q.

Proof of Theorem 4. The normal distribution belongs to the exponential family and hence $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma})$ is an analytic function of $\underline{\theta}$ for each Σ. Thus $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma}) = \alpha$ if $\underline{\theta}^{(1)} = \pm\Delta$ due to the unbiasedness of Φ. Let A be the set of $\underline{\theta}$ such that $|\underline{\theta}^{(1)}| < \Delta$. Since, by assumption, $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma})$ is unimodal in $\underline{\theta}^{(1)}$, it is not smaller than α on A. On the other hand $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma})$ should be no larger than α on $A \cap H_0$ because of its unbiasedness. This implies $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma})$ is constant α on $A \cap H_0$ which contains an open set of $\underline{\theta}$. Therefore, by analyticity, $E_{\underline{\theta},\Sigma}\Phi(\bar{X},\hat{\Sigma}) = \alpha$ for all $\underline{\theta}$ and Σ. We conclude $\Phi(\bar{X},\hat{\Sigma}) = \alpha$ almost surely by completeness and the proof of the theorem is complete.

References

[1] Berger, R.L. (1982), "Multiparameter Hypothesis Testing and Acceptance Sampling," *Technometrics*, 24, 295-300.

[2] Berger, R.L. and Hsu, J. (1996), "Bioequivalence Trials, Intersection-Union Tests, and Equivalence Confidence Sets," to appear in *Statistical Sciences.*

[3] Brown, L.D., Casella, G. and Hwang, J.T.G, (1995), "Optimal Confidence Sets, Bioequivalence, and the limacon of Pascal," *Journal of the American Statistical Association*, 90, 880-889

[4] Brown, L.D., Hwang, J.T.G, and Munk, A. (1996), " An Unbiased Test for the Bioequivalence problem," to appear in *Annals of Statistics.*

[5] Casella, G. and Berger, R.L. (1990), *Statistical Inference*, Duxbury Press.

[6] FDA (1992), " FDA Guidance on Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design," *Division of Bioequivalence, Office of Generic Drugs, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Rockville, MD.*

[7] Hsu, H. C., Lu, H.L. and Chan, K.K.H. (1995), " A Novel Multivariate Approach for Estimating the Degree of Similarity in Bioavailability between Two Pharmaceutical Products," *Journal of Pharmaceutical Sciences,* 84, 768-772.

[8] Schall R., Groenewald, P.C.N., Potgieter, L.A. and Luus, H.G. (1996), "Multivariate Assessment of Bioequivalence," *Technical Report, University of the Orange Free State, R.S.A..*

[9] Schuirmann, D.J. (1987), "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability," *Journal of Pharmacokinetics and Biopharmaceutics,* 15, 657-680.

[10] Wang, W. (1995), *On Assessment of Bioequivalence*, Ph. D. thesis, Cornell University.