Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 138 (2008) 3982–3992
journal of statistical planning and inference
www.elsevier.com/locate/jspi

A p-value for testing the equivalence of the variances of a bivariate normal distribution

Thomas Mathew[a,*], Gitanjali Paul[a,b]

[a]*Department of Mathematics and Statistics, University of Maryland, 1000 Hilltop Circle, Baltimore, MD 21250, USA*
[b]*GlaxoSmithKline, King of Prussia, Pennsylvania, PA 19406, USA*

Received 18 May 2005; received in revised form 20 April 2007; accepted 19 February 2008
Available online 29 February 2008

Abstract

A p-value is developed for testing the equivalence of the variances of a bivariate normal distribution. The unknown correlation coefficient is a nuisance parameter in the problem. If the correlation is known, the proposed p-value provides an exact test. For large samples, the p-value can be computed by replacing the unknown correlation by the sample correlation, and the resulting test is quite satisfactory. For small samples, it is proposed to compute the p-value by replacing the unknown correlation by a scalar multiple of the sample correlation. However, a single scalar is not satisfactory, and it is proposed to use different scalars depending on the magnitude of the sample correlation coefficient. In order to implement this approach, tables are obtained providing sub-intervals for the sample correlation coefficient, and the scalars to be used if the sample correlation coefficient belongs to a particular sub-interval. Once such tables are available, the proposed p-value is quite easy to compute since it has an explicit analytic expression. Numerical results on the type I error probability and power are reported on the performance of such a test, and the proposed p-value test is also compared to another test based on a rejection region. The results are illustrated with two examples: an example dealing with the comparability of two measuring devices, and an example dealing with the assessment of bioequivalence.

MSC: 62F03; 62H15; 62P99

Keywords: Bioequivalence; Crossover design; Incomplete beta; Wishart distribution

1. Introduction

We address the problem of testing if the variances of a bivariate normal distribution are equivalent, i.e., if they are close according to a specified criterion. Let $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$ be the variance–covariance matrix of a bivariate normal distribution. The hypothesis of interest to us is the equivalence of σ_{11} and σ_{22}, i.e., we want to test if the ratio σ_{11}/σ_{22} is close to 1. The hypothesis can be stated as

$$\mathrm{H}_0 : \frac{\sigma_{11}}{\sigma_{22}} \geqslant c \quad \text{or} \quad \frac{\sigma_{11}}{\sigma_{22}} \leqslant \frac{1}{c} \quad \text{vs.} \quad \mathrm{H}_1 : \frac{1}{c} < \frac{\sigma_{11}}{\sigma_{22}} < c, \tag{1}$$

* Corresponding author. Tel.: +1 410 455 2418; fax: +1 410 455 1066.
E-mail address: mathew@math.umbc.edu (T. Mathew).

0378-3758/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2008.02.005

for a suitably chosen $c > 1$. Note that we conclude the equivalence of σ_{11} and σ_{22} if H_0 is rejected. Define the parameter

$$\theta = \max\left(\frac{\sigma_{11}}{\sigma_{22}}, \frac{\sigma_{22}}{\sigma_{11}}\right). \tag{2}$$

Then our hypotheses can equivalently be stated as

$$H_0 : \theta \geqslant c \quad \text{vs.} \quad H_1 : \theta < c. \tag{3}$$

An article that addresses the above testing problem is Wang (1999a), who has derived a rejection region for the problem. Wang's (1999a) test is quite satisfactory; it is not too conservative. Our purpose here is to derive an easily computable p-value.

The above problem is of interest in several applications. The problem is relevant in the context of establishing the equivalency of measuring devices; for example, when we want to establish the equivalence of an alternative measuring device to a standard device. Wellek (2002, p. 85) reports an example dealing with the comparability of blood pressure measurements taken using two different automatic devices. The sample for this consists of bivariate observations on the diastolic blood pressure of 20 individuals obtained using the two devices (the data are reproduced in Table 6). Note that even if the means are equivalent, we cannot conclude that the two devices are equivalent, unless the variances are also equivalent. The problem of comparing two measuring devices is very common in applications in industrial hygiene, where it is required to obtain data on workplace exposure to toxicants. For gathering such data, industrial hygienists would prefer to use a cheap or easy to use sampling device, provided it is equivalent to an accurate standard device. For background information and examples on this, we refer to the article by Krishnamoorthy and Mathew (2002).

The testing problem (1) is also of interest in the assessment of bioequivalence in two settings. Suppose a 2×2 crossover design is used, with n subjects, to test the bioequivalence of a test drug T with a reference drug R. Then each subject receives T and R once, and let Y_{jT} and Y_{jR} denote the corresponding responses for the jth subject. Typically, the response obtained is the area under the curve (AUC), or the maximum blood concentration ($C_{\max}$) after a log transformation. Several authors have used the one-way random model for Y_{jT} and Y_{jR}; see, for example, Sheiner (1992), Schall and Luus (1993), Schall (1995) and Wang (1999b). Thus the model is

$$\begin{aligned} Y_{jT} &= \mu_T + \eta_{jT} + \varepsilon_{jT}, \\ Y_{jR} &= \mu_R + \eta_{jR} + \varepsilon_{jR}, \end{aligned} \tag{4}$$

$j = 1, 2, \ldots, n$, where μ_T and μ_R are population mean responses corresponding to treatments T and R, η_{jT} and η_{jR} are random subject effects, and ε_{jT} and ε_{jR} are the random within-subject errors. It is further assumed that (η_{jT}, η_{jR}) follows a bivariate normal distribution with zero means and variance–covariance matrix, say Σ_B, given by

$$\Sigma_B = \begin{pmatrix} \sigma^2_{BT} & \sigma_{BTR} \\ \sigma_{BTR} & \sigma^2_{BR} \end{pmatrix}, \tag{5}$$

and $\varepsilon_{jT} \sim N(0, \sigma^2_{WT})$ and $\varepsilon_{jR} \sim N(0, \sigma^2_{WR})$. Define

$$\bar{Y}_T = \frac{1}{n}\sum_{j=1}^{n} Y_{jT}, \quad \bar{Y}_R = \frac{1}{n}\sum_{j=1}^{n} Y_{jR},$$

$$S = \sum_{j=1}^{n} \begin{pmatrix} Y_{jT} - \bar{Y}_T \\ Y_{jR} - \bar{Y}_R \end{pmatrix} (Y_{jT} - \bar{Y}_T, Y_{jR} - \bar{Y}_R),$$

$$\Sigma = \Sigma_B + \operatorname{diag}(\sigma^2_{WT}, \sigma^2_{WR}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \quad \text{(say)}.$$

Then it is easily verified that $\hat{\mu}_T = \bar{Y}_T$ and $\hat{\mu}_R = \bar{Y}_R$ are unbiased estimators of μ_T and μ_R. Furthermore,

$$\operatorname{Var}\begin{pmatrix} Y_{jT} \\ Y_{jR} \end{pmatrix} = \Sigma, \quad \operatorname{Var}\begin{pmatrix} \hat{\mu}_T \\ \hat{\mu}_R \end{pmatrix} = \frac{1}{n}\Sigma \quad \text{and} \quad S \sim W_2(\Sigma, n-1),$$

where $W_2(\Sigma, n-1)$ denotes the two-dimensional Wishart distribution with $n-1$ degrees of freedom and associated variance–covariance matrix Σ. We note that the alternative hypothesis in (1), or equivalently in (3), states that $\hat{\mu}_T$ and $\hat{\mu}_R$ (and also Y_{jT} and Y_{jR}) have equivalent variances. This is also noted in Wang (1999a).