Statistics and Probability Letters 117 (2016) 118–127

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The finite sample performance of the two-stage analysis of a two-period crossover trial

Paul Kabaila *

Department of Mathematics and Statistics, La Trobe University, Victoria 3086, Australia

ARTICLE INFO

Article history:
Received 28 March 2016
Received in revised form 17 May 2016
Accepted 17 May 2016
Available online 25 May 2016

Keywords:
Crossover trial
Differential carryover
Large-sample approximation
Preliminary hypothesis test
Two-stage procedure

ABSTRACT

For subject and error variances assumed known, Freeman assesses the two-stage analysis of an AB/BA crossover trial. We provide a finite sample assessment of this analysis for the practical situation that these variances are estimated from the data.

1. Introduction

We consider a two-treatment two-period (AB/BA) crossover trial, with responses that are continuous random variables. This design is very popular in a wide range of medical and other applications, see e.g. Jones and Kenward (2003), Senn (2002) and Senn (2006). The purpose of this trial is to carry out inference about the difference θ in the effects of two treatments, labeled A and B. Subjects are randomly allocated to either group 1 or group 2. Subjects in group 1 receive treatment A in the first period and then receive treatment B in the second period. Subjects in group 2 receive treatment B in the first period and then receive treatment A in the second period. This design is efficient under the assumption that there is no differential carryover effect. It is not an appropriate design unless there is strong prior information that this assumption holds.

However, a commonly occurring scenario is that it is not certain that the assumption of no differential carryover holds. We consider this scenario. To deal with this uncertainty, it has been suggested (starting with Grizzle, 1965, 1974 and endorsed by Hills and Armitage, 1979) that the following two-stage procedure be used. In the first stage, we carry out a preliminary test of the null hypothesis that this assumption holds. In the second stage, the result of this first stage is used as follows. If this test leads to acceptance of this null hypothesis then further inference proceeds on the basis that it was known *a priori* that there is no differential carryover effect. If, on the other hand, this null hypothesis is rejected then further inference is based solely on data from the first period, since this is unaffected by any carryover effect.

In a landmark paper, Freeman (1989) considered this two-stage procedure for the construction of a confidence interval for the treatment difference θ, with nominal coverage $1-\alpha$. In this case, the second stage is as follows. If the null hypothesis of no differential carryover is accepted then the confidence interval with coverage $1-\alpha$ is used, assuming that it was known *a priori* that the differential carryover is zero. If, on the other hand, this null hypothesis is rejected then the confidence interval with coverage $1-\alpha$ and based solely on the data from the first period is used. He shows that the confidence interval resulting

* Tel.: +61 03 9479 2594; fax: +61 03 9479 2466.
E-mail address: P.Kabaila@latrobe.edu.au.

http://dx.doi.org/10.1016/j.spl.2016.05.011
0167-7152/

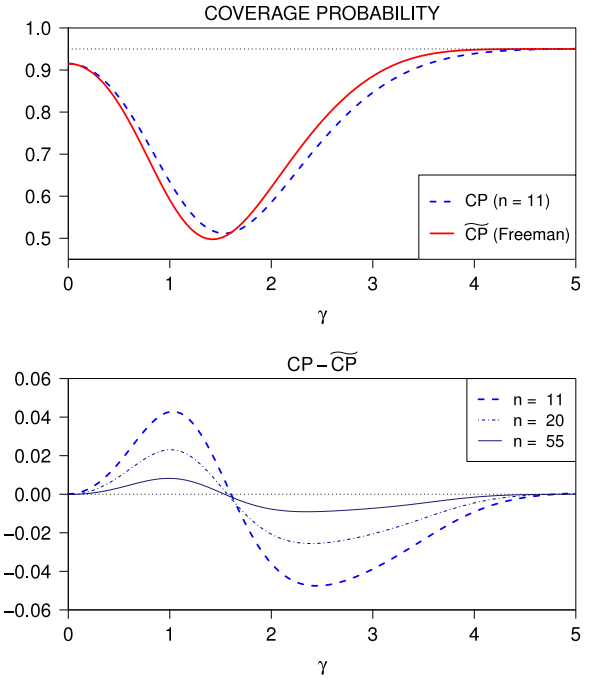

Fig. 1. The top panel consists of a graph of $\widetilde{CP}$ (coverage probability found by Freeman (1989), assuming the subject and error variances are known) and a graph of CP (coverage probability when subject and error variances are estimated from the data) for $n = 11$, as functions of γ, a scaled version of the differential carryover. The lower panel consists of graphs of $CP - \widetilde{CP}$ as functions of γ for $n = 11$, 20 and 55. Both panels are for a level of significance 0.1 of the preliminary test of the null hypothesis of no differential carryover, confidence interval nominal coverage probability 0.95 and (subject variance)/(subject variance + error variance) equal to 0.6.

from this two-stage procedure has minimum coverage probability far below $1 - \alpha$. Freeman's conclusion that the use of a preliminary test in this way 'is too potentially misleading to be of practical use' is now widely accepted (Senn, 2006).

For simplicity, Freeman supposes that the subject variance and the error variance are known. Kabaila and Vicendese (2012) make similar assumptions in their assessment of the performance of a two-stage analysis of ABAB/BABA crossover trials. Of course, in practice, the subject and the error variances are unknown and must be estimated from the data. However, it is plausible that Freeman's assumption that these variances are known corresponds to a large-sample analysis, when these variances are estimated from the data.

We suppose that, as must be the case in practice, both the subject and error variances are estimated from the data. We rely on the work of Grieve (1987) who shows that if the treatment difference θ is estimated solely from the data for period 1 then finding a confidence interval for this difference is equivalent to the Behrens–Fisher problem. We use a frequentist ('sampling theory') evaluation of the two-stage analysis and so we have chosen to use Welch's 'approximate degrees of freedom' (also called the 'approximate t') solution to the Behrens–Fisher problem (Welch, 1949, p. 295).

Let CP denote the coverage probability, given by our new Theorem 1, of the confidence interval for θ resulting from the two-stage procedure, with nominal coverage $1 - \alpha$, when the subject and error variances are estimated from the data. Also let $\widetilde{CP}$ denote the coverage probability, derived by Freeman (1989), of this confidence interval when the subject and error variances are assumed to be known. We measure the sample size by $n = n_1 + n_2 - 2$, where n_1 and n_2 are the numbers of subjects in groups 1 and 2, respectively.

The main purpose of the present paper is to show that, for a given level of significance of the preliminary hypothesis test and given nominal coverage $1 - \alpha$, CP approaches $\widetilde{CP}$, as n approaches infinity. This is shown numerically in Fig. 1, for a level of significance 0.1 of the preliminary test of the null hypothesis of no differential carryover, confidence interval nominal coverage probability 0.95 and (subject variance)/(subject variance + error variance) equal to 0.6. These are the same values as those considered in Fig. 2 of Freeman (1989). The top panel of Fig. 1 consists of a graph of $\widetilde{CP}$ and a graph of CP for $n = 11$, as functions of γ, a scaled version of the differential carryover. The lower panel of Fig. 1 consists of graphs of $CP - \widetilde{CP}$ as functions of γ for $n = 11$, 20 and 55. As n increases, $CP - \widetilde{CP}$ approaches zero.

Our new Theorem 2 proves that this is true for all possible values of preliminary test level of significance, confidence interval nominal coverage probability and (subject variance)/(subject variance + error variance), not just those particular values considered in Fig. 1. Fig. 1 shows numerically that Freeman's expression for the coverage probability $\widetilde{CP}$ provides an excellent approximation to the finite sample coverage probability CP, even for moderate sample sizes n. In other words, we have confirmed Freeman's conclusion that the confidence interval for θ that results from the two-stage procedure should not be used.