

Accepted Manuscript

A class of tests for the two-sample problem for count data

M.V. Alba-Fernández, A. Batsidis, M.D. Jiménez-Gamero, P. Jodrá

PII: S0377-0427(16)30479-4

DOI: <http://dx.doi.org/10.1016/j.cam.2016.09.050>

Reference: CAM 10838

To appear in: *Journal of Computational and Applied Mathematics*

Received date: 14 July 2016

Revised date: 22 September 2016



Please cite this article as: M.V. Alba-Fernández, A. Batsidis, M.D. Jiménez-Gamero, P. Jodrá, A class of tests for the two-sample problem for count data, *Journal of Computational and Applied Mathematics* (2016), <http://dx.doi.org/10.1016/j.cam.2016.09.050>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A class of tests for the two-sample problem for count data

M.V. Alba-Fernández^{a,*}, A. Batsidis^b, M.D. Jiménez-Gamero^c, P. Jodrá^d

^a*Department of Statistics and O.R., University of Jaén, Spain*

^b*Department of Mathematics, University of Ioannina, Greece*

^c*Department of Statistics and O.R., University of Seville, Spain*

^d*Department of Statistical Methods, University of Zaragoza, Spain*

Abstract

A class of tests for the two-sample problem for count data whose test statistic is an L_2 -norm of the difference between the empirical probability generating functions associated with each sample is considered. The tests can be applied to count data of any arbitrary fixed dimension. Since the null distribution of the test statistic is unknown, some approximations are investigated. Specifically, the bootstrap, permutation and weighted bootstrap estimators are examined. All of them provide consistent estimators. A simulation study analyzes the performance of these approximations for small and moderate sample sizes. This study also includes a comparison with other two-sample tests whose test statistic is a weighted integral of the difference between the empirical characteristic functions of the samples.

Keywords: two-sample problem, count data, probability generating function, simulation

1. Introduction

The two-sample problem, which consists on testing whether two samples come from the same population, is a statistical issue of great interest and many different approaches have been proposed to deal with it (see, for example, Baringhaus and Kolbe [1] for a recent paper on this topic and the references therein). One of them is related to the use of the characteristic function (CF) and its empirical counterpart (ECF) by means of an L_2 -norm between the ECFs associated with each sample. This kind of tests can be applied to all sort of data, continuous, discrete or mixed of any arbitrary fixed dimension. Since the resultant test statistic is not

*Corresponding author

Email addresses: mvalba@ujaen.es (M.V. Alba-Fernández), abatsidis@uoi.gr (A. Batsidis), dolores@us.es (M.D. Jiménez-Gamero), pjodra@unizar.es (P. Jodrá)

Download English Version:

<https://daneshyari.com/en/article/5776424>

Download Persian Version:

<https://daneshyari.com/article/5776424>

[Daneshyari.com](https://daneshyari.com)