Revista Mexicana de Ciencias Geológicas
ISSN: 1026-8774
rmcg@geociencias.unam.mx
Universidad Nacional Autónoma de México
México

Critical values for six Dixon tests for outliers in normal samples up to sizes 100, and applications in science and engineering

Surendra P. Verma[1,2,*] and Alfredo Quiroz-Ruiz[1]

[1] *Centro de Investigación en Energía, Universidad Nacional Autónoma de México, Priv. Xochicalco s/no., Col Centro, Apartado Postal 34, Temixco 62580, Mexico.*
[2] *Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001, Col. Chamilpa, Cuernavaca 62210, Mexico*
* *spv@cie.unam.mx*

ABSTRACT

In this paper we report the simulation procedure along with new, precise, and accurate critical values or percentage points (with 4 decimal places; standard error of the mean ≤0.0001) for six Dixon discordance tests with significance levels α = 0.30, 0.20, 0.10, 0.05, 0.02, 0.01, 0.005 and for normal samples of sizes n up to 100. Prior to our work, critical values (with 3 decimal places) were available only for n up to 30, which limited the application of Dixon tests in many scientific and engineering fields. With these new tables of more precise and accurate critical values, the applicability of these discordance tests (N7 and N9-N13) is now extended to 100 observations of a particular variable in a statistical sample. We give examples of applications in many diverse fields of science and engineering including geosciences, which illustrate the advantage of the availability of these new critical values for a wider application of these six discordance tests. Statistically more reliable applications in science and engineering to a greater number of cases can now be achieved with our new tables than was possible earlier. Thus, we envision that these new critical values will result in wider applications of the Dixon tests in a variety of scientific and engineering fields such as agriculture, astronomy, biology, biomedicine, biotechnology, chemistry, environmental and pollution research, food science and technology, geochemistry, geochronology, isotope geology, meteorology, nuclear science, paleontology, petroleum research, quality assurance and assessment programs, soil science, structural geology, water research, and zoology.

Key Words: Outlier methods, normal sample, Monte Carlo simulations, reference materials, earth sciences.

RESUMEN

En este trabajo se presenta el procedimiento para la simulación junto con valores críticos o puntos porcentuales nuevos y más precisos y exactos (con 4 puntos decimales; el error estándar de la media ≤0.0001) de las seis pruebas de discordancia de Dixon y para los niveles de significancia α = 0.30, 0.20, 0.10, 0.05, 0.02, 0.01, 0.005 y para tamaños n de las muestras normales de hasta 100. Antes de nuestro trabajo, se disponía de valores críticos (con 3 puntos decimales) solamente para n hasta 30, lo cual limitaba seriamente la aplicación de las pruebas de Dixon en muchos campos de las ciencias e ingenierías. Con las nuevas tablas de valores críticos más precisos y exactos obtenidos en el presente trabajo, la aplicabilidad de las pruebas de Dixon (N7 y N9-N13) se ha extendido a 100 observaciones de una variable en una muestra estadística. Presentamos ejemplos de aplicaciones en muchos campos de ciencias e ingenierías incluyendo las geociencias. Estos ejemplos demuestran la ventaja de la disponibilidad de estos nuevos *valores críticos para una aplicación muy amplia de esas seis pruebas de discordancia. Se esperan aplicaciones a un mayor número de casos en ciencias e ingenierías, estadísticamente más confiables que como era posible anteriormente. De esta manera, prevemos que los nuevos valores críticos resulten en aplicaciones de las pruebas de Dixon mucho más amplias en una variedad de campos de ciencias e ingenierías tales como agronomía, astronomía, biología, biomedicina, biotecnología, ciencia del suelo, ciencia nuclear, ciencia y tecnología de los alimentos, contaminación ambiental, geocronología, geología estructural, geología isotópica, geoquímica, investigación del agua y del petróleo, programas de aseguramiento y evaluación de calidad, paleontología, química, meteorología y zoología.*

Palabras clave: Métodos de valores desviados, muestra normal, simulaciones Monte Carlo, materiales de referencia, pruebas de discordancia de Dixon, Ciencias de la Tierra.

INTRODUCTION

Two main sets of methods (Outlier methods and Robust methods; Barnett and Lewis, 1994) exist for correctly estimating location (central tendency) and scale (dispersion) parameters for a set of experimental data likely to be drawn, in most cases in science and engineering, from a normal or Gaussian distribution (Verma, 2005). The outlier scheme is based on a set of tests for normality (or detection of outliers) such as Dixon tests described here. However, caution is required when applying such outlier tests for samples that are not normally distributed. The alternative scheme for arriving at these parameters consists of a series of robust or accommodation approach methods (for location parameter: *e.g.*, median, mode, Winsorized mean, trimmed mean, and mean quartile; and for scale parameter: *e.g.*, interquartile range and median deviation; see Barnett and Lewis, 1994; Verma, 2005, or any standard text book on statistics), all of which rely on not "taking into account" the outlying and other peripheral observations in a set of experimental data. These methods, although in use in many branches of science and engineering, will not be considered here any further because the main objective of this paper is to comment on and improve the applicability of six discordance tests, proposed by Dixon more than 50 years ago, which are still widely used as explained below.

Dixon (1950, 1951, 1953) proposed six discordance tests for normal univariate samples and estimated critical values or percentage points for these tests for sizes up to 30 and reported them to 3 decimal places. These tests were designated N7 and N9-N13 by Barnett and Lewis (1994). Dixon (1951) also stated that the estimated critical values for tests N7 (test statistic r_{10} in this paper), N9 (statistic r_{11}), and N10 (statistic r_{12}) were "in error by not more than one or two units in the third (decimal) place", whereas those for tests N11 (statistic r_{20}), N12 (statistic r_{21}), and N13 (statistic r_{22}) were "believed to be accurate to within three or four units in the third (decimal) place".

These tests have been widely used –and are still in use– in the outlier-based scheme for correctly estimating the location and scale parameters (*e.g.*, Thomulka and Lange, 1996; Freeman *et al.*, 1997; Hanson *et al.*, 1998; Verma *et al.*, 1998; Woitge *et al.*, 1998; Muranaka, 1999; Tigges *et al.*, 1999; Taylor, 2000; Hofer and Murphy, 2000; Buckley and Georgianna, 2001; Langton *et al.*, 2002; Reed *et al.*, 2002; Stancak *et al.*, 2002; Yurewicz, 2004; Kern *et al.*, 2005). However, these tests are applicable to only samples of sizes up to 30, which severely limits their application in many scientific and engineering fields, because, today, the number of individual data in a statistical sample has considerably increased (to much greater than 30) than was customary a few decades ago. Furthermore, Gawlowski *et al.* (1998) considered the Dixon tests for normal univariate samples as inferior to the Grubbs tests because the critical values for the former (quoted to only three significant digits, or 3 decimal places; Dixon, 1951) are less accurate than for the latter (quoted to four significant digits, or 3 or 4 decimal places depending on the critical values being >1 or <1; Grubbs and Beck, 1972). In fact, other reasons (see pp. 121-125 and p. 222 in Barnett and Lewis, 1994) might account for the relative efficiency of discordance tests than the one stated by Gawlowski *et al.* (1998).

The computation of new critical values for Dixon discordance tests through Monte Carlo simulations was motivated from multiple reasons: (1) The still wide use of these tests by researchers in many scientific and engineering fields (see selected references for the past ten years 1996-2005 cited above); (2) the availability of critical values for Dixon tests with 3 decimal places as compared to Grubbs tests with critical values with 3 or 4 decimal places; and most importantly (3) the inapplicability of these discordance tests to the actual data for numerous chemical elements in reference materials (RMs) in the field of (a) alloy industry (*e.g.*, Roelandts, 1994); (b) biology (Ihnat, 2000); (c) biomedicine (Patriarca *et al.*, 2005); (d) cement industry (Sieber *et al.*, 2002); (e) food industry (In't Veld, 1998, Langton *et al.*, 2002); (f) environmental research (Dybczyński *et al.*, 1998; Gill *et al.*, 2004; Holcombe *et al.*, 2004); (g) rock geochemistry (*e.g.*, Guevara *et al.*, 2001); and (h) soil science (Dybczyński *et al.*, 1979; Hanson *et al.*, 1998; Verma *et al.*, 1998), as well as to experimental data in numerous other scientific and engineering applications