Agrociencia
ISSN: 1405-3195
agrocien@colpos.mx
Colegio de Postgraduados
México

Villaseñor Alva, José A.; Díaz Carreño, Miguel A.
Pruebas no paramétricas para procesos poisson no homogéneos
Agrociencia, vol. 37, núm. 1, enero-febrero, 2003, pp. 21-31
Colegio de Postgraduados
Texcoco, México

PRUEBAS NO PARAMÉTRICAS PARA PROCESOS POISSON NO HOMOGÉNEOS

NON-PARAMETRIC TESTS FOR NON-HOMOGENEOUS POISSON PROCESSES

José A. **Villaseñor-Alva** y Miguel A. **Díaz-Carreño**

Especialidad de Postgrado en Estadística. Instituto de Socioeconomía, Estadística e Informática. Colegio de Postgraduados. 56230. Montecillo, Estado de México. (jvillasr@colpos.mx) (mdiaz@colpos.mx)

Resumen

El análisis de varios tipos de fenómenos aleatorios que ocurren en el tiempo, en ocasiones se realiza bajo el supuesto de un proceso estocástico Poisson no homogéneo (PPNH), sin la validación previa de este supuesto. Esta investigación propone un método estadístico para probar este supuesto, método que se basa en una caracterización de los PPNH. El método considera la utilización de las estadísticas de Cramér-von Mises (CVM), como una prueba de bondad de ajuste para la distribución Poisson, y la prueba de Kruskal-Wallis (KW). Las pruebas basadas en las estadísticas de CVM presentan mayor potencia que otras pruebas de bondad de ajuste comúnmente empleadas para la distribución Poisson. Aquí se presenta un estudio por simulación, el cual muestra que la prueba de KW tiene mayor potencia que la de Cramér-von Mises para evaluar si dos o más muestras provienen de una misma distribución, contra algunas alternativas de interés.

Palabras clave: Estadísticas de Cramér-von Mises, prueba de Kruskal-Wallis, pruebas de bondad de ajuste, procesos Poisson.

Abstract

The analysis of several types of random phenomena that occur in time is on occasion carried out under the assumption of a non-homogeneous Poisson stochastic process (NHPP), without previous validation. This paper proposes a statistical method for testing this assumption, which is based on a characterization of the NHPP. The method considers the use of Cramér-von Mises statistics (CVM), as a goodness of fit test for the Poisson distribution and the Kruskal-Wallis (KW) test. The tests based on CVM statistics are more powerful than other goodness of fit tests commonly used for the Poisson distribution. Here a simulation study is presented, which shows that the KW test has more power than that of the Cramér-von Mises test to evaluate whether or not two or more samples have the same distribution against some alternatives of interest.

Key Words: Cramér-von Mises' statistics, test of Kruskal-Wallis, goodness of fit tests, Poisson processes.

Introducción

En diversas áreas del conocimiento como la economía, la botánica y la epidemiología, se generan gran cantidad de fenómenos aleatorios como el número de alzas o descensos en la bolsa de valores más allá de un nivel extremo con repercuciones prácticas, el número de plantas con alguna plaga por área determinada, etc.

Estos fenómenos, en ocasiones, se modelan como procesos Poisson no homogéneos (PPNH), aun cuando pudieran ser de otra naturaleza.

El objetivo de esta investigación consiste en desarrollar una metodología estadística, a partir de una caracterización de los PPNH, que permita probar si una serie de eventos puede modelarse como un PPNH.

Marco Teórico

El proceso estocástico $\{N(t), t\geq 0\}$, donde $N(t)$ es el número de eventos que ocurrieron hasta el tiempo t, es llamado un PPNH con función de intensidad $\lambda(t)$ si:

Recibido: Febrero, 2001. Aprobado: Diciembre, 2002.
Publicado como ENSAYO en Agrociencia 37: 21-31. 2003.

Introduction

In several areas of knowledge, such as economics, botany, and epidemiology, among others, a large number of random phenomena occur, for example, the number of rises and falls in the stock market beyond an extreme level with practical repercussions, the number of plants with some kind of pest in a given area, etc.

These phenomena are sometimes modeled as non-homogeneous Poisson processes (NHPP), even when they might be of some other nature.

The objective of this study was to develop a statistical method, based on a characterization of the NHPP, that would allow testing to determine whether a series of events can be modeled as a NHPP.

Theoretical Framework

The stochastic process $\{N(t), t\geq 0\}$, where $N(t)$ is the number of events that occurred up to time t, is called a NHPP with an intensity function $\lambda(t)$ if:

i) $N(0) = 0$
ii) $\{N(t), t\geq 0\}$ has independent increases, or if the

i) $N(0)=0$
ii) $\{N(t), t\geq 0\}$ tiene incrementos independientes; o bien, si el número de eventos que ocurren en intervalos de tiempo ajenos son independientes.
iii) $P\{N(t+s)-N(t)=n\}=e^{-[\Lambda(t+s)-\Lambda(t)]}.$

$$\frac{[\Lambda(t+s)-\Lambda(t)]^n}{n!} \text{ con } n\geq 0 \text{ (Ross, 1980).}$$

Entonces, $N(t+s)-N(t)$ tiene una distribución Poisson con media $\Lambda(t+s)-\Lambda(t)$ y $N(t)$ se distribuye Poisson con media $\Lambda(t)$, por lo cual $\Lambda(t)$ es llamada función de valor medio del proceso y se define por $\Lambda(t)=\int_0^t \lambda(s)ds$. Si $\lambda(t)=\eta$, entonces $\Lambda(t)=\eta t$ y se tendrá un proceso Poisson homogéneo (PPH) con media ηs (Kao, 1997).

Caracterización del proceso Poisson no homogéneo (PPNH)

Chouinard y McDonald (1985) presentan una caracterización del PPNH con base en la distribución de los intervalos de los tiempos de arribo.

Sean $\{N_1(t), N_2(t),...,N_q(t); 0\leq t\leq T\}$ q repeticiones independientes del proceso $N(t)$, en donde $N_1(T)=n_1, N_2(T)=n_2,..., N_q(T)=n_q$, con n_i = número de eventos en la repetición i, i=1,2,...,q, donde $n=n_1+n_2+...+n_q$. Además, sean $\{S_{i(j)}\}_{j=1}^{n_i}$ los tiempos en que ocurren los eventos o tiempos de arribos y defínase a $R_{i(j)}$, i=1,...,q, j=1,...,n_i como el rango del j-ésimo tiempo de arribo del proceso $\{N_i(t); 0\leq t\leq T\}$ entre los n tiempos de arribos.

Proposición 1. Si $N(t)$ es un PPNH, entonces para toda q y todo conjunto $\{n_i\}_{i=1}^{q}$

$$P\{R_{i(j)}=k_{i(j)}; i=1,\ldots,q, j=1,\ldots,n_i / N_i(T)=n_i; i=1,\ldots,q\} = \left(\frac{n!}{n_1!\,n_2!,\ldots,n_q!}\right)^{-1}$$

donde $\{k_{i(j)}\}_{j=1}^{n_i}$ es una secuencia creciente de enteros distintos para cada i, $\{k_{i(j)};\ j=1,...,n_i, i=1,...,q\}=\{1,2,..,n\}$. Nótese que los rangos tienen distribución uniforme discreta.

Teorema 1. N(t) es un PPNH si y sólo si:

a) la distribución marginal de N(T) es Poisson y
b) N(t) satisface la proposición 1.

number of events that occur in disjoint time intervals are independent.

iii) $P\{N(t+s)-N(t)=n\}=e^{-[\Lambda(t+s)-\Lambda(t)]}.$

$$\frac{[\Lambda(t+s)-\Lambda(t)]^n}{n!} \text{ with } n\geq 0 \text{ (Ross, 1980).}$$

Then, $N(t+s)-N(t)$ has a Poisson distribution with mean $\Lambda(t+s)-\Lambda(t)$, and $N(t)$ is Poisson distributed with mean $\Lambda(t)$; thus $\Lambda(t)$ is called a mean value function of the process and is defined by $\Lambda(t)=\int_0^t \lambda(s)ds$. If $\lambda(t)=\eta$, then $\Lambda(t)=\eta t$, and a homogeneous Poisson process will result with mean ηs (Kao, 1997).

Characterization of the non-homogeneous Poisson process (NHPP)

Chouinard and McDonald (1985) present a characterization of the NHPP based on the distribution of the ranges of arrival times.

Let $\{N_1(t), N_2(t),...,N_q(t); 0\leq t\leq T\}$ be q independent repetitions of the $N(t)$ process, where $N_1(T)=n_1$, $N_2(T)=n_2$,..., $N_q(T)=n_q$, with n_i = number of events in the repetition i, i=1, 2, …, q, where $n=n_1+n_2+...+n_q$. Also, let $\{S_{i(j)}\}_{j=1}^{n_i}$ be the times at which the events occur, or arrival times, and define $R_{i(j)}$, i=1,...,q, j=1,...,n_i as the rank of the j^{th} arrival time of the process $\{N_i(t); 0\leq t\leq T\}$ between the n arrival times.

Proposition 1. If $N(t)$ is a NHPP, then for all q and all sets $\{n_i\}_{i=1}^{q}$

$$P\{R_{i(j)}=k_{i(j)}; i=1,\ldots,q, j=1,\ldots,n_i / N_i(T)=n_i; i=1,\ldots,q\} = \left(\frac{n!}{n_1!\,n_2!,\ldots,n_q!}\right)^{-1}$$

where $\{k_{i(j)}\}_{j=1}^{n_i}$ is an ascending sequence of different integers for each i, $\{k_{i(j)};\ j=1,...,n_i, i=1,...,q\}=\{1,2,..,n\}$. Note that the ranges have a uniform discrete distribution.

Theorem 1. $N(t)$ is a NHPP if, and only if:

a) the marginal distribution of $N(t)$ is Poisson, and
b) $N(t)$ satisfies proposition 1.

Based on this characterization, a statistical method to test for NHPP must take into account the ranks of the observations.

Con base en esta caracterización, un método estadístico de prueba para PPNH debe tomar en cuenta los rangos de las observaciones.

Nótese que la validación del supuesto de PPNH para un conjunto de eventos dependerá del cumplimiento de a) y b) en el teorema 1. La prueba de la primera condición se debe basar en una prueba de bondad de ajuste para la distribución Poisson, y la de la segunda, en una prueba no paramétrica para igualdad de funciones de distribución.

Prueba de bondad de ajuste para la distribución Poisson

La revisión de literatura para pruebas de bondad de ajuste para la distribución Poisson reveló que la prueba basada en las estadísticas de Cramér-von Mises (CVM) tiene mayor potencia contra distribuciones alternativas como la Binomial, Beta-binomial, Binomial negativa y Uniforme discreta, todas con parámetros definidos (Spinelli y Stephens, 1997). Por lo que esta prueba es una de las mejores opciones para evaluar la primera condición del Teorema 1.

Estadísticas de Cramér-von Mises

En esta sección se describe una prueba de bondad de ajuste basada en una de las estadísticas de CVM comúnmente usada para probar la hipótesis nula: las variables n_i, i=1,2,...,q son una muestra aleatoria de alguna distribución discreta especificada.

Las variables aleatorias n_i pueden tomar cualquier valor entero no negativo. Sea p_j la probabilidad de que cualquiera de las variables aleatorias n_i tome el valor j. Supóngase además que se tienen q observaciones independientes de n_i: $n_1, n_2,...,n_q$. Sea o_j el número observado o frecuencia de las observaciones iguales a j, y $qp_j=e_j$ el número esperado de valores j, $W_j = \sum_{i=0}^{j}(o_i - e_i)$ y $H_j = \sum_{i=0}^{j} p_i$, donde $i, j = 0,1,2,....$ La estadística de prueba es

$$A^2 = q^{-1}\sum_{J=0}^{\infty}\frac{W_j^2 p_j}{H_j\left(1-H_j\right)} \tag{1}$$

Cuando los parámetros de la distribución son desconocidos, éstos deberán ser estimados con la muestra. Sea $\hat{p}_j$ la probabilidad estimada del valor observado j, y $\hat{e}_j$, $\hat{W}_j$ y $\hat{H}_j$ calculados reemplazando p_j por $\hat{p}_j$. La esta-

Note that the validation of the NHPP assumption for a set of events will depend on fulfilling a) and b) in Theorem 1. The test for the first condition must be based on a goodness of fit test for the Poisson distribution, and the second on a non-parametric test for equality of distribution functions.

Goodness of fit test for the Poisson distribution

Revision of literature on goodness of fit tests for Poisson distribution revealed that the test based on the Cramér-von Mises (CVM) statistics are more powerful when compared against alternatives such as the Binomial, Beta-Binomial, Negative Binomial, and Discrete Uniform, all of which have defined parameters (Spinelli and Stephens, 1997). This test thus constitutes one of the best options for evaluating the first condition of Theorem 1.

Cramer-von Mises statistics

In this section a goodness of fit test based on CVM statistics commonly used to test the null hypothesis is described: The variables n_i, i=1,2,...,q are a random sample of some specified discrete distribution.

The random variables n_i can take any non-negative integer value. Let p_j be the probability that any of the random variables n_i takes the value j. Assume, besides, that there are q independent observation of n_i: $n_1, n_2,...,n_q$. Let o_j be the observed number, or frequency, of the observations equal to j, and $qp_j=e_j$ the expected number of values j, $W_j = \sum_{i=0}^{j}(o_i - e_i)$ and $H_j = \sum_{i=0}^{j} p_i$, where i,j=0,1,2,.... The test statistic is:

$$A^2 = q^{-1}\sum_{J=0}^{\infty}\frac{W_j^2 p_j}{H_j\left(1-H_j\right)} \tag{1}$$

When the parameters of the distribution are unknown, they will have to be estimated with the sample. Let $\hat{p}_j$ be the estimated probability of the observed value j, and $\hat{e}_j$, $\hat{W}_j$ and $\hat{H}_j$ calculated replacing p_j by $\hat{p}_j$. The statistic (1) will be calculated using $\hat{p}_j$, $\hat{W}_j$ and $\hat{H}_j$. Also, in practice, the sum should be finite, according to the available information.

For the Poisson distribution $p_j = \frac{\mu^j e^{-\mu}}{j!}$, $j = 0,1,2,...$, the following procedure allows testing the H_0 hypothesis.

dística (1) se calculará usando $\hat{p}_j$, $\hat{W}_j$ y $\hat{H}_j$. Además, en la práctica, la suma deberá ser finita, según la información disponible.

Para la distribución Poisson, $p_j = \frac{\mu^j e^{-\mu}}{j!}$, $j = 0,1,2,\ldots$, el siguiente procedimiento permite probar la hipótesis H_0.

1) Calcular la estadística A^2, estimando a μ por $\overline{X} = q^{-1}\sum_{i=1}^{q} n_i$.
2) Comparar el valor calculado de A^2 con un valor de tablas de acuerdo con μ estimada (Spinelli y Stephens, 1997).
3) Regla de decisión: rechazar H_0 a un nivel de significancia α cuando el valor de A^2 sea mayor al de tablas.

Pruebas estadísticas no paramétricas para igualdad de poblaciones

Enseguida se presentan dos pruebas estadísticas no paramétricas comúnmente empleadas para probar igualdad de poblaciones.

Prueba de Kruskal-Wallis (KW)

Sean $S_{11}, S_{12},..., S_{1n_1}$; $S_{21}, S_{22},...,...$; S_{11}, S_{12} ,..., S_{qn_q}, q muestras aleatorias de tamaño n_i con i=1,2,...,q que provienen de funciones de distribución desconocidas F_1, F_2, ..., F_q. Se quiere probar la hipótesis nula:

H_0: $F_1(s) = F_2(s) = ... = F_q(s) = F(S)$ contra la alternativa H_1: $F_i(s) = F(s-\theta_i)$ (para toda s, i=1,2,...,q), donde las θ_i son números reales no necesariamente iguales.

La estadística de prueba es

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{q}\frac{R_i^2}{n_i} - 3(n+1) \qquad (2)$$

donde $R_i = \sum_{j=1}^{n_i} R_{ij}$ es la suma de los rangos de los n_i valores de la i-ésima muestra y H se distribuye aproximadamente como Ji-cuadrada con $q-1$ grados de libertad $\left(\chi^2_{q-1}\right)$ cuando $n_i \geq 5$ (Kruskal, 1952). Puesto que valores pequeños de H apoyan H_0, entonces se rechaza H_0 cuando $H > \chi^2_{\alpha,q-1}$.

1) To calculate the statistic A^2, estimating μ by $\overline{X} = q^{-1}\sum_{i=1}^{q} n_i$.
2) To compare the calculated value of A^2 with a value from tables, according to the estimated μ (Spinelli and Stephens, 1997).
3) Decision rule: reject H_0 at a selected α level when the value of A^2 is greater than that of the tables.

Non-parametric statistics tests for equality of populations

In this section, two non-parametric statistics tests commonly used to test equality of populations are presented.

Kruskal-Wallis (KW) test

Let $S_{11}, S_{12},..., S_{1n_1}$; $S_{21}, S_{22},..., S_{2n_2}$; ...; $S_{11}, S_{12},..., S_{qn_q}$, be q random samples of size n_i with i=1, 2, ..., q, from the unknown distribution functions F_1, F_2,...,F_q. We are interested in testing the null hypothesis: H_0: $F_1(s) = F_2(s) = ... = F_q(s) = F(S)$ against the alternative H_1: $F_i(s) = F(s-\theta_i)$ (for all s, i=1, 2, …, q), where the θ_i's are real numbers not necessarily equal.

The test statistic is

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{q}\frac{R_i^2}{n_i} - 3(n+1) \qquad (2)$$

where $R_i = \sum_{j=1}^{n_i} R_{ij}$ is the sum of the ranks of the n_i values of the i^{th} sample, and H is distributed approximately as chi-square with $q-1$ degrees of freedom $\left(\chi^2_{q-1}\right)$ when $n_i \geq 5$ (Kruskal, 1952). Because small values of H support H_0, H_0 is rejected when $H > \chi^2_{\alpha,q-1}$.

Cramer-von Mises (CVM) Test

Chouinard and McDonald (1985) discuss the CVM non-parametric test considering a problem of q samples with n_i observations in the i sample (arrival times in $N_i(t)$). The hypothesis of interest is:

H_0: $F_1(s) = F_2(s) = ... = F_q(s)$ against
H_1: At least one of the q samples come from a different population.

Prueba de Cramer-von Mises (CVM)

Chouinard y McDonald (1985) discuten la prueba no paramétrica de CVM considerando un problema de q muestras con n_i observaciones en la muestra i (tiempos de arribos en $N_i(t)$). La hipótesis de interés es:

H_0: $F_1(s) = F_2(s) = ... = F_q(s)$ contra
H_1: Al menos una de las q muestras proviene de una población distinta.

Ahora, sea F_{n_i} la función de distribución empírica de la i-ésima muestra y $\overline{F}$ la distribución empírica de las $n = \sum_{i=1}^{q} n_i$ observaciones. La estadística de prueba es

$$\Re = \sum_{i=1}^{q} \int_{-\infty}^{\infty} n\left(\frac{n_i}{n_i+1}F_i(t) - \frac{n}{n+1}\overline{F}(t)\right)^2 dF_i(t) \sum_{i=1}^{q}\sum_{j=1}^{n_i}\left(\frac{R_{i(j)}}{n+1}\frac{j}{n_i+1}\right)^2$$

donde $\Re$ mide la variación entre las distribuciones muestrales y $R_{i(j)}$ es el rango de la j-ésima observación en la i-ésima muestra.

Bajo la hipótesis de distribuciones muestrales idénticas, $\Re$ se puede expresar de la siguiente manera

$$\Re = \frac{n(2n+1)}{6(n+1)} + \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)} - 2\sum_{i=1}^{q}\sum_{j=1}^{n_i}\left(\frac{j}{(n_i+1)}\frac{R_{i(j)}}{n+1}\right)$$

El valor esperado y la varianza de $\Re$ están dados por

$$E(\Re) = \frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}$$

$$\mathrm{Var}(\Re) = \frac{1}{45(n+1)}\left\{n(q-1) + (2q-2n-5)\left(\sum_{i=1}^{q}\frac{1}{n_i-1}\right)\right.$$

$$\left. +(n+2)\left(\sum_{i=1}^{q}\frac{1}{(n_i-1)^2}\right) - \left(\sum_{i=1}^{q}\frac{1}{n_i-1}\right)^2 + q(-q+3)\right\}$$

Por el Teorema Central del Límite (TCL) $\lim_{q\to\infty}\frac{\Re - E(\Re)}{\sqrt{Var\Re}} = Z \sim N(0,1)$. Entonces, una prueba no paramétrica para medir variabilidad distribucional entre poblaciones puede basarse en $\Re$.

Now, let F_{n_i} be the empirical distribution of the i^{th} sample and $\overline{F}$ the empirical distribution of all of the $n = \sum_{i=1}^{q} n_i$ observations. The test statistic is

$$\Re = \sum_{i=1}^{q} \int_{-\infty}^{\infty} n\left(\frac{n_i}{n_i+1}F_i(t) - \frac{n}{n+1}\overline{F}(t)\right)^2 dF_i(t) \sum_{i=1}^{q}\sum_{j=1}^{n_i}\left(\frac{R_{i(j)}}{n+1}\frac{j}{n_i+1}\right)^2$$

where $\Re$ measures the variation among the sample distributions, and $R_{i(j)}$ is the rank of the j^{th} observation in the i^{th} sample.

Under the hypothesis of identical sample distributions, $\Re$ can be expressed in the following way

$$\Re = \frac{n(2n+1)}{6(n+1)} + \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)} - 2\sum_{i=1}^{q}\sum_{j=1}^{n_i}\left(\frac{j}{(n_i+1)}\frac{R_{i(j)}}{n+1}\right)$$

The expected value and the variance of $\Re$ are given by

$$E(\Re) = \frac{n(2n+1)}{6(n+1)} - \sum_{i=1}^{q}\frac{n_i(2n_i+1)}{6(n_i+1)}$$

$$\mathrm{Var}(\Re) = \frac{1}{45(n+1)}\left\{n(q-1) + (2q-2n-5)\left(\sum_{i=1}^{q}\frac{1}{n_i-1}\right)\right.$$

$$\left. +(n+2)\left(\sum_{i=1}^{q}\frac{1}{(n_i-1)^2}\right) - \left(\sum_{i=1}^{q}\frac{1}{n_i-1}\right)^2 + q(-q+3)\right\}$$

Because of the Central Limit Theorem of (CLT) $\lim_{q\to\infty}\frac{\Re - E(\Re)}{\sqrt{Var\Re}} = Z \sim N(0,1)$. Hence, a non-parametric test to measure the distributional variability among populations can be based on $\Re$.

Because small values of $\Re$ support the assumption that the q samples come from identical populations, the test rejects H_0 when $\Re$ takes large values.

So, we are interested in $P(\Re > k / H_0) \leq \alpha$, where k is a critical value for a given level of significance α, determined with base on the asymptotic distribution of the $\Re$ statistic. Then, $k_\alpha = \left(\sqrt{\mathrm{Var}\,\Re}\right)\Phi^{-1}(1-\alpha) + E(\Re)$, Dado que valores pequeños de $\Re$ apoyan el supuesto de que las q muestras provienen de poblaciones idénticas, la prueba rechaza H_0 cuando $\Re$ toma valores grandes.

Por tanto nos interesa que $P(\Re > k / H_0) \leq \alpha$, donde k es un valor crítico para cierto nivel de significancia α, determinado con base en la distribución asintótica de la estadística $\Re$. Entonces, $k_\alpha = \left(\sqrt{\mathrm{Var}\,\Re}\right)\Phi^{-1}(1-\alpha) + E(\Re)$, donde Var($\Re$) y E($\Re$) pueden ser sustituidos por sus respectivas estimaciones muestrales S_q^2 y $\overline{\Re}_q$, además $\Phi(y)$ es la función de distribución normal estándar.

Por lo tanto, si $\Re > k_\alpha$ se rechaza la hipótesis H_0 al nivel de significancia α.

Estudio del tamaño y potencia de las pruebas de KW y CVM

En esta sección se presenta un estudio por simulación del tamaño y potencia de las pruebas de KW y CVM para la evaluación de la segunda condición que debe satisfacer un PPNH.

Estudio de simulación para el tamaño de las pruebas

El tamaño de una prueba estadística se denota por α y se define teóricamente como la probabilidad máxima de rechazar H_0 cuando es verdadera.

En el estudio del tamaño de una prueba estadística por simulación, lo que se obtiene es una estimación de α.

La estimación del tamaño de las pruebas en estudio se interpreta como la probabilidad de rechazar el modelo de PPNH, dado que la distribución marginal de N(T) es Poisson, cuando en realidad este es el modelo adecuado para el conjunto de datos que se está analizando.

La estimación del tamaño de la prueba de KW se basa en el Algoritmo(1).

Algoritmo(1). 1) Simular q {q = 5, 8, 10, 12, 15, 20, 25, 30, 50} números aleatorios de la distribución Poisson con media 30. Cada q indica el número de muestras o repeticiones del proceso $N(t)$, y la media, el número de arribos u observaciones por muestra.

2) Simular n datos aleatorios (tiempos de arribos) de un PPNH del tipo Weibull con función de valor medio $\Lambda(t)=3t^2$ (López *et al*., 2002) tal que n_1 datos pertenecen a la primera muestra, n_2 a la segunda, y así sucesivamente, hasta n_q que es el número de datos en la q-ésima muestra. ($n=n_1+n_2+...+n_q$). La generación de los n tiempos de arribos se basa en la ecuación $S= [10^*U]^{1/2}$, donde U es una variable aleatoria uniforme en el intervalo (0,3.16).

where Var($\Re$) and E($\Re$) can be substituted by their respective sample estimations S_q^2 and $\overline{\Re}_q$. Also, $\Phi(y)$ is the standard normal distribution function.

Therefore, if $\Re > k_\alpha$ the null hypothesis is rejected at significance level α.

Study of size and power of the KW and CVM tests

This section presents a study by simulation of the size and power of the KW and CVM tests for the evaluation of the second condition that a NHPP must satisfy.

Simulation study for test size

The size of a statistical test is denoted by α and is defined theoretically as the maximum probability of rejecting H_0 when it is true.

In the study of the size of a statistical test by simulation, an estimation of α is obtained.

The estimation of the size of the tests in the study is interpreted as the probability of rejecting the NHPP model, given that the marginal distribution of $N(T)$ is Poisson, when it is actually the suitable model for the set of data that is being analyzed.

The estimation of the size of the KW test is based on Algorithm(1).

Algorithm(1). 1) Simulate q {q = 5, 8, 10, 12, 15, 20, 25, 30, 50}, random numbers of the Poisson distribution with mean 30. Each q indicates the number of samples or replicates of the $N(t)$ process, and the mean indicates the number of arrivals or observations per sample.

2) Simulate n random data (arrival times) of a Weibul-type NHPP with mean value function $\Lambda(t)=3t^2$ (López *et al*., 2002), such that n_1 data belong to the first sample, n_2 to the second, and so forth, up to n_q, which is the number of data in the q^{th} sample. $n = (n_1 + n_2 + ... + n_q)$. Generation of n arrival times is based on the equation $S= [10^*U]^{1/2}$, where U is a random variable uniform in the interval (0, 3.16).

3) Calculate the statistic H of KW given by equation (2) with the q samples and compare it with a percentile of the distribution $\chi^2_{\alpha, q-1}$, α={0.01, 0.05, 0.10, 0.20, 0.25}. If $H > \chi^2_{\alpha, q-1}$, H_0 is rejected.

4) Repeat the procedure 1000 times and the number of rejections of H_0 divided by 1000 estimates the KW test size.

The results are shown in Figure 1, where it observed that for different q values the probability of type I error is less or equal to the level of significance α. Thus the KW test maintains its size in this case.

3) Se calcula la estadística H de KW dada por la ecuación (2) con las q muestras y se compara con un percentil de la distribución $\chi^2_{\alpha,q-1}$, α={0.01, 0.05, 0.10, 0.20, 0.25}. Si $H > \chi^2_{\alpha,q-1}$, se rechaza H_0.

4) Repetir el procedimiento 1000 veces y el número de rechazos de H_0 entre 1000 estima el tamaño de la prueba de KW.

Los resultados se muestran en la Figura 1, donde se observa que para diferentes valores de q la probabilidad del error tipo I resulta menor o igual que el nivel de significancia α; por lo que la prueba de KW preserva su tamaño en este caso.

La estimación del tamaño de la prueba de CVM se basa en el Algoritmo(2).

Algoritmo(2). 1) Realizar los pasos 1 y 2 del Algoritmo(1).

2) Calcular $\sqrt{S_q^2}$ y $\overline{\Re}_2$ para las q muestras.

3) Calcular la estadística $\Re$ de CVM con las q muestras y compararla con el valor k_α para un nivel de significancia α {0.01, 0.05, 0.10, 0.20, 0.25}. Si $\Re > k_\alpha$ se rechaza H_0.

4) Este procedimiento se repite 1000 veces y el número de rechazos de H_0 entre 1000 estima el tamaño de la prueba de CVM.

En la Figura 2 se presentan los tamaños estimados de prueba. La prueba de CVM preserva su tamaño α en este ejemplo, puesto que para los distintos valores de q, la probabilidad del error tipo I es menor que el nivel de significancia α.

En ambos procedimientos se supone un PPNH del tipo Weibull, por lo que las pruebas de KW y de CVM no rechazan H_0 excepto en el nivel α especificado.

The estimation of the CVM test size is based on Algorithm(2).

Algorithm(2). 1) Carry out steps 1 and 2 of Algorithm(1).

2) Calculate $\sqrt{S_q^2}$ and $\overline{\Re}_2$ for the q samples.

3) Calculate the CVM statistic $\Re$ with the q samples and compare it with the k_α value for the level of significance α {0.01,0.05,0.10,0.20,0.25}. If $\Re > k_\alpha$, H_0 is rejected.

4) This procedure is repeated 1000 times, and the number of rejections of H_0 divided by 1000 estimates the CVM test size.

Figure 2 presents the estimated test sizes. The CVM test preserves its α size in this example since for the different values of q, the probability of type I error is lower than the level of significance α.

In both procedures a Weibull-type NHPP is assumed, and so the KW and CVM tests do not reject H_0, except at a specified level of α.

Simulation study for power of the tests

Power (δ) of a statistical test is defined as

$$\delta = P \text{ (reject } H_0 \text{ when it is false)}$$

In the estimation of power for both tests, two alternative cases are considered.

First, it is assumed that in H_1 we have a process of renovation with times between arrivals with a normal multivariate distribution $N(\underline{\mu}, \sum)$ of n_i dimension with a means vector $\underline{\mu} = (30,\ldots,30)_{1xn_i}$, where n_i~$P(30)$;

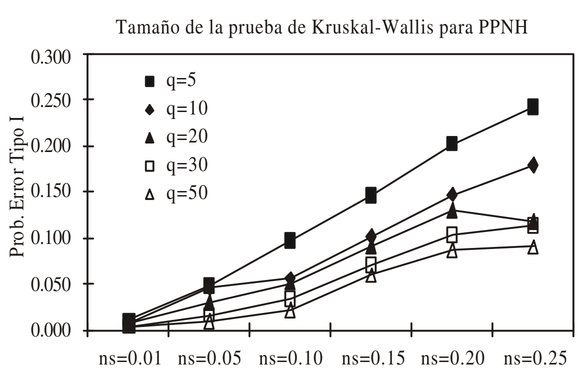

Figura 1. Probabilidad estimada del error tipo I de la prueba cuando las muestras provienen de un PPNH del tipo Weibull. ns=nivel de significancia, q=número de muestras.

Figure 1. Estimated probability of type I error of the test when the samples come from a Weibull-type NHPP. ns=level of significance, q=number of samples.

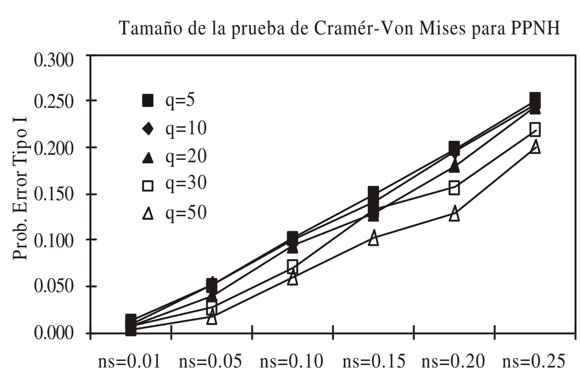

Figura 2. Probabilidad estimada del error tipo I de la prueba cuando las muestras provienen de un PPNH del tipo Weibull. ns=nivel de significancia, q=número de muestras.

Figure 2. Estimated probability of type I error of the test when the samples come from a Weibull-type NHPP. ns=level of significance, q=number of samples.

Estudio de simulación para la potencia de las pruebas

La potencia (δ) de una prueba estadística se define como

$$\delta = P \text{ (rechazar } H_0 \text{ cuando es falsa)}$$

En la estimación de la potencia para ambas pruebas se consideran dos casos alternativos.

Primero se supone que se tiene en H_1 un proceso de renovación con tiempos entre llegadas que tienen una distribución normal multivariada $N(\underline{\mu}, \Sigma)$ de dimensión n_i, con vector de medias $\underline{\mu} = (30,\ldots,30)_{1xn_i}$, donde n_i~$P(30)$; i =1, 2,...,q y q = 5, 8, 10, 15, 20, 25, 30, 50, y con matriz de varianzas y covarianzas

$$\Sigma = \begin{pmatrix} 1 & 1/2 & \ldots & 1/2 \\ 1/2 & 1 & \ldots & 1/2 \\ \vdots & \vdots & \vdots & \vdots \\ 1/2 & 1/2 & 1/2 & 1 \end{pmatrix}_{n_i x n_i} \quad \text{(Mardia, } et\ al.\text{, 1979).}$$

En la segunda alternativa se considera un proceso cuyos tiempos de arribo son generados por la serie de tiempo $\tau_i = \alpha\tau_{i-1} + \beta\tau_{i-2}$, con α = 0.5 y β = 0.3 (Hanke y Reitsh, 1996).

Puesto que la distribución conjunta de los tiempos de arribo $\{S_1, S_2,...,S_n\}$ de un PPNH, dado que $N(T) = n$ es la misma que la conjunta de las estadísticas de orden de n variables aleatorias independientes e idénticamente distribuidas en el intervalo $[0,T]$ (McDonald,1994), y considerando además que los tiempos entre arribos de un PPNH son variables aleatorias independientes de alguna distribución continua (Ross, 1980), nótese que en el primer caso los tiempos entre arribos no son variables aleatorias independientes, por lo que ese proceso no es un PPNH.

Dado que las estadísticas de orden de una muestra aleatoria de variables continuas tienen una estructura de dependencia Markoviana (Barry *et al.*, 1992), un proceso estocástico generado por una serie de tiempo con al menos dos rezagos no tiene esta propiedad, por lo que el proceso del segundo caso será distinto al PPNH.

La estimación de la potencia de la prueba de KW cuando en H_1 se tiene el proceso de renovación antes descrito, se realizó por medio del Algoritmo(3).

Algoritmo(3). 1) Se realiza el paso 1 del Algoritmo(1).

2) Se simulan n_i variables aleatorias de la distribución normal multivariada $N(\underline{\mu}, \Sigma)$ de dimensión n_i, especificada anteriormente. Enseguida, utilizando las n_i variables (tiempos entre llegadas) se generan los tiempos

i=1, 2,...,q and q = 5, 8, 10, 15, 20, 25, 30, 50, and with matrix of variances and co-variances

$$\Sigma = \begin{pmatrix} 1 & 1/2 & \ldots & 1/2 \\ 1/2 & 1 & \ldots & 1/2 \\ \vdots & \vdots & \vdots & \vdots \\ 1/2 & 1/2 & 1/2 & 1 \end{pmatrix}_{n_i x n_i} \quad \text{(Mardia, } et\ al.\text{, 1979).}$$

The second alternative considers a process whose arrival times are generated by the time series $\tau_i = \alpha\tau_{i-1} + \beta\tau_{i-2}$, with α = 0.5 and β = 0.3 (Hanke and Reitsh, 1996).

Since the joint distribution of arrival times $\{S_1, S_2,...,S_n\}$ of a NHPP, given that $N(T)=n$ is the same as the joint distribution of the order statistics of n variables that are independent, random, and identically distributed in the interval $[0,T]$ (McDonald,1994), and considering also that times between arrivals of a NHPP are independent random variables of some continuous distribution (Ross, 1980), note that in the first case the times between arrivals are not independent random variables, and for that reason the process is not a NHPP.

Given that the order statistics of a random sample of continuous variables have a structure of Markovian dependence (Barry, *et al.*, 1992), a statistical process generated by a time series with at least two delays does not have this property, and so the process in the second case will be different from a NHPP.

Estimation of the power of the KW test, when the process of renovation described above exists in H_1, was done by means of Algorithm(3).

Algorithm(3). 1) Carry out step 1 of Algorithm(1).

2) Simulate n_i random variables of the multivariate normal distribution $N(\underline{\mu}, \Sigma)$ of dimension n_i, specified previously. Then, using the n_i variables (arrival times), the arrival times of sample i, i=1, 2,…, q, with the equation $S_{n_i} = X_1 + X_2 + ... + X_{n_i}$, $n_i \geq 1$. This procedure is repeated q times, to have q replicates of $N(t)$.

3) Carry out steps 3 and 4 of Algorithm(1).

Estimation of the power of the CVM test when H_1 asumes the renovation process is calculated with Algorithm(4).

Algorithm(4). 1) Carry out steps 1 and 2 of Algorithm (3).

2) Solve steps 2, 3, and 4 of Algorithm(2).

In Figure 3 it can be observed that the KW test has high power when the process of renovation is assumed in H_1. It approaches 0.80 if q=10 with a level of significance of 0.10, while with q=20, power is 0.90 for a level of significance of 0.10. Under the same conditions, the CVM test has a power of 0.58 for 10 samples and 0.64 for 20.

de llegada de la muestra i, i=1,2,...,q mediante la ecuación $S_{n_i} = X_1 + X_2 + ... + X_{n_i}$, $n_i \geq 1$. Este procedimiento se repite q veces, para tener q réplicas de $N(t)$.

3) Se llevan al cabo los pasos 3 y 4 del Algoritmo(1).

La estimación de la potencia de la prueba de CVM cuando en H_1 se supone el proceso de renovación, se calcula con el Algoritmo(4).

Algoritmo(4). 1) Realizar los pasos 1 y 2 del Algoritmo(3).

2) Desarrollar los pasos 2, 3 y 4 del Algoritmo(2).

En la Figura 3 se puede observar que la prueba de KW presenta una potencia elevada cuando en H_1 se asume el proceso de renovación; ésta se acerca a 0.80 si q=10 con un nivel de significancia de 0.10, mientras que con q=20 la potencia es de 0.90 para un nivel de significancia de 0.10. Bajo las mismas condiciones la prueba de CVM tiene potencia de 0.58 para 10 muestras y de 0.64 para 20.

La estimación de la potencia de la prueba de KW cuando en H_1 se tiene la serie de tiempo antes descrita, se realizó por medio del Algoritmo(5).

Algoritmo(5). 1) Se realiza el paso 1 del Algoritmo(1).

2) Se simulan n datos aleatorios con la ecuación $\tau_i = \alpha\tau_{i-1} + \beta\tau_{i-2}$, con α = 0.5 y β = 0.3, donde τ_{i-1} y τ_{i-2} son generados por la distribución Uniforme (0,2). Además, n_1 datos pertenecen a la primera muestra, n_2 a la segunda, y así sucesivamente hasta n_q que es el número de datos en la q-ésima muestra ($n = n_1+n_2+...+n_q$).

3) Se llevan a cabo los pasos 3 y 4 del Algoritmo(1).

Por otra parte, la potencia de la prueba de CVM, cuando en H_1 se tiene el proceso generado por la serie de tiempo, se estimó con el Algoritmo(6).

Algoritmo(6). 1) Realizar los pasos 1 y 2 del Algoritmo(5).

2) Desarrollar los pasos 2, 3 y 4 del Algoritmo(2).

En la Figura 4 se muestra la potencia de las dos pruebas cuando en H_1 se supone la serie de tiempo arriba especificada. Nótese que para la prueba de KW cuando q=10 y con nivel de significancia 0.05, la potencia es superior a 0.90. Además, cuando q=20 la potencia de la prueba se aproxima a 1.00 para un nivel de significancia de 0.05 o mayor. Bajo las mismas condiciones, la prueba de CVM tiene una potencia de 0.65 para q=10 y de 0.72 con 20 muestras o réplicas de $N(t)$.

Este estudio de simulación muestra evidencia en favor de la prueba de KW, la cual presenta mayor potencia que la prueba de CVM cuando se evalúa si un proceso estocástico contador de eventos es un PPNH, bajo el supuesto de que la distribución marginal de $N(T)$ es Poisson. Las Figuras 3 y 4 apoyan esta apreciación.

Con base en el estudio comparativo realizado en esta sección, el método estadístico de prueba que se propone para la validación del supuesto de un PPNH incluye dos

Estimation of the power of the KW test when in H_1 is considered the time series described above was done with Algorithm(5).

Algorithm(5). 1) Carry out step 1 of Algorithm(1).

2) Simulate n random data with the equation $\tau_i = \alpha\tau_{i-1} + \beta\tau_{i-2}$, with α = 0.5 and β = 0.3, where τ_{i-1} and τ_{i-2} are generated by the Uniform distribution (0,2). Also, n_1 data belong to the first sample, n_2 to the second, and so on, up to n_q, which is the number of data in the q^{th} sample, ($n = n_1+n_2+...+n_q$).

3) Carry out steps 3 and 4 of Algorithm (1).

Power of the CVM test was estimated with Algorithm(6) when H_1 includes the process generated by the time series.

Algorithm(6). 1) Do steps 1 and 2 of Algorithm(5).

2) Solve steps 2, 3, and 4 of Algorithm(2).

Figure 4 shows the power of the two tests when the specified time series is assumed in H_1. Note that for the KW test, when q=10 and with a significance level of 0.05, the power is higher than 0.90. When q=20 the power approaches 1.00 for a significance level of 0.05 or higher. Under the same conditions, CMV test has a power of 0.65 for q=10 and of 0.72 with 20 samples or replicates of $N(t)$.

This simulation study shows evidence in favor of the KW test, which is more powerful than the CMV test when evaluating if a counting events stochastic process is a NHPP, under the assumption that the marginal distribution of $N(T)$ is Poisson. Figures 3 and 4 support this appreciation.

Based on the comparative study conducted in this section, the statistical test method proposed for the validation of the assumption of a NHPP, considers two

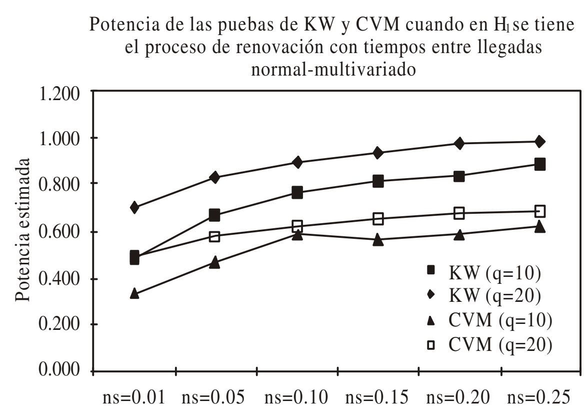

Figura 3. Potencia estimada de las pruebas de KW y CVM cuando en H_1 se supone el proceso de renovación; ns= nivel de significancia, q=número de muestras.

Figure 3. Estimated power of the KW and CVM tests when the process of renovation is assumed in H_1; ns=level of significance, q=number of samples.

etapas: en la primera se aplica la prueba de bondad de ajuste basada en la estadística A^2 de CVM y, en la segunda, se usa la prueba no paramétrica de KW.

Una aplicación

A continuación se presenta una aplicación del método propuesto a un conjunto de datos relacionados con los tiempos de fallas de una computadora (Lewis, 1964) (Cuadro 1).

Al realizar las pruebas correspondientes, se obtuvieron los siguientes resultados.

El valor calculado de la estadística A^2 (1.0364) es menor al de tablas (1.187) para un nivel de significancia de 0.05, dado que μ=120.67. Por lo tanto, no se rechaza la hipótesis de que la distribución marginal de estos datos es Poisson al nivel de significancia indicado; además, el valor calculado de la estadística de KW es H = 3.02, el cual es menor que el percentil $\chi^2_{0.05,\,2} = 5.99$, por lo que la hipótesis nula no se rechaza al nivel de significancia de 0.05. Se concluye que los tiempos de fallas de una computadora pueden ser modelados como un PPNH.

Conclusiones

El método de prueba propuesto está basado en una caracterización de los PPNH. De acuerdo con el estudio de simulación realizado en este trabajo, se concluye que para un tamaño de prueba especificado (0.05) la prueba de KW tiene mayor potencia que la prueba de CVM para la validación de un PPNH con respecto a las alternativas analizadas, bajo el supuesto de que las distribuciones marginales del proceso son Poisson.

Los resultados obtenidos de la aplicación del método de prueba al conjunto de datos de Lewis (1964) son consistentes con lo expuesto por López (2002).

Literatura citada

Barry, C. A., N. Balakrisnhan, and H. N. Nagaraja. 1992. A first course in order statistics. John Wiley and Sons, New York, N.Y. 279 p.

Chouinard A. and McDonald D. 1985. A Characterization of Non-Homogeneous Poisson Processes. Stochastics, 15: 113-119.

Hanke, J. E., and Reitsh, A. G. 1996. Business forecasting. 5th edition. Prentice Hall Inc., New York, N.Y. 605 p.

Kao, E. P. C. 1997. An introduction to stochastic processes. Wadsworth. Belmont, California. 438 p.

Kruskal, W. H. 1952. A nonparametric test for the several sample Problem. The Annals of Mathematical Statistics, 23(4): 525-540.

Lewis, P. A. W. 1964. A Branching Poisson Processes Model for the Analysis of computer Failure Patterns. J. Royal Statist. Soc. B, 26: 398-456.

López, S. L, Villaseñor, A. J., y Vaquera H. 2002. Dos pruebas de bondad de ajuste para procesos Poisson no homogéneos. Agrociencia, Vol. 36(6). pp: 703-712.

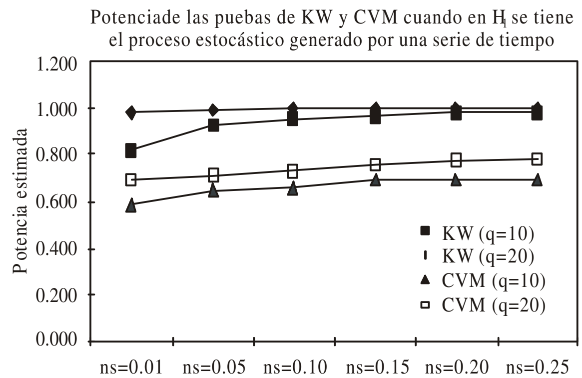

Figura 4. Potencia estimada de las pruebas de KW y de CVM cuando en H_1 se supone un proceso generado por la serie de tiempo $\tau_i = \alpha\tau_{i\text{-}1} + \beta\tau_{i\text{-}2}$. ns= nivel de significancia, q=número de réplicas de N(t).

Figure 4. Estimated power of the KW and CVM tests when a process generated by the time series $\tau_i = \alpha\tau_{i\text{-}1} + \beta\tau_{i\text{-}2}$ is assumed in H_1. ns= level of significance, q=number of replicates of N(t).

stages: in the first, the goodness of fit test based on the statistic A^2 of CVM is applied and, in the second, the non-parametric KW test is used.

An application

In this section the proposed method is applied to a set of data related to times of computer failures (Lewis, 1964) (Table 1).

After performing the corresponding tests, the following results were obtained:

The calculated value of the A^2 statistic (1.0364) is smaller than that of tables (1.187) for a level of significance of 0.05, given that μ=120.67. Therefore, the hypothesis that the marginal distribution of these data is Poisson is not rejected at the indicated level of significance. Also, the calculated value of the KW statistic is H=3.02, which is smaller than the percentile $\chi^2_{0.05,\,2}$ = 5.99, and thus the null hypothesis is not rejected at the level of significance of 0.05. It is concluded that the failure times of a computer can be modeled as a NHPP (See Table 1 of Appendices).

Conclusions

The test method proposed is based on a characterization of NHPP. According to the simulation study conducted in this research it is concluded that for a specified size (0.05), the KW test is more powerful than the CVM test for the validation of a NHPP, compared with the alternatives analyzed, under the assumption that the marginal distributions of the process are Poisson.

Cuadro 1. Tiempos de falla de una computadora, tres repeticiones.
Table 1. Computer failure times, three replicates.

				Repetición 1					
943	1190	2288	2504	2681	3279	3761	3899	4499	4985
5084	5097	5197	5339	5496	6369	8216	8262	8346	8769
9096	9106	9614	9617	9665	9682	9742	9810	9950	10185
10977	12090	12108	12336	12369	13021	13296	14144	14153	14244
14249	14324	14329	14663	16732	16861	17091	17319	18579	18721
18963	18989	19071	19081	19133	19401	19473	19659	19704	20033
20037	20647	20649	20743	21013	21127	21171	21302	21823	21879
22540	22969	23834	24469	24495	25330	25348	25778	26029	26042
26334	26346	26485	26525	27537	27782	27797	27850	28590	28792
28855	28867	29027	29674	29709	30732	30920	30930	30966	31210
31523	33568	34062	34104	34728	34788	35034	35080	35183	
				Repetición 2					
9	66	92	132	149	289	621	721	799	814
850	876	946	987	1498	1680	2203	2218	2266	2291
2826	3884	4126	4219	4226	4247	4267	4378	4427	4438
4618	4915	5046	5093	5583	5758	5943	7331	7341	7384
7389	7997	8158	8257	8704	8714	8949	9064	9274	9306
9570	9799	9804	9939	9959	9993	10630	11888	12934	13180
13244	13621	12778	13883	14679	15809	15965	16360	16453	16459
16818	16888	17006	17504	17542	18118	18880	18890	19337	19352
19679	19744	20071	20650	20820	20825	21707	21920	22320	22681
22770	23135	23540	23575	23625	23640	23793	23808	23811	23916
23918	24047	24703	24850	25162	25496	25907	26359	26814	26889
26894	27408	27441	27534	27538	27635	27698	27711	27834	28065
28364	28409	28632	30345	30365	30383	30395	30785	31119	31630
31722	32007	32037	32182	32207	32861	33081	33599	33604	33617
33657	33894	34069	34239	34359	34447				
				Repetición 3					
340	570	604	1358	1469	1484	1619	1764	1862	2155
2420	2566	2800	2888	3001	3289	3335	3408	3643	3714
4490	4508	4868	5161	5199	5228	5745	6273	6841	7863
9147	9387	11642	11955	12003	12435	12682	13333	13612	13774
13948	14281	14620	14676	15030	15033	15166	15170	16864	16999
17939	17950	18133	18147	18279	18525	18533	18554	18870	18875
19056	19286	19567	20327	20328	20645	20722	20942	20991	21176
21709	22004	22481	22780	22910	23040	23073	23080	23241	23253
23348	23439	23504	23662	23732	23787	25854	26816	27230	27461
27889	27891	28427	28745	28779	29303	29442	29461	29478	29489
29592	29614	30011	30019	30035	31166				

Mardia, K. V., J. T. Kent, and J. M. Bibby, 1979. Multivariate analysis. Academic Press, New York, N.Y. 518 p.
McDonald, D. 1994. Elements of applied probability for engineering mathematics and systems science. University of Ottawa, Ontario, Canadá. 311 p.
Ross, S. M. 1980. Introduction to probability models. 2nd edition. Academic Press. New York, N.Y. 478 p.
Spinelli, J. J. and Stephens, M. A. 1997. Cramér-von Mises tests of fit for the Poisson distribution. The Canadian Journal of Statistics, 25(2): 257-268.

The results obtained from the application of the test method to Lewis' (1964) set of data are consistent with that presented by López (2002).

—End of the English version—