



**ISSN: 1561-3194**

***Rev. Ciencias Médicas. octubre 2010; 14(4):***

**BIOESTADÍSTICA**

## **Estadística Piloto: paquete estadístico digital educacional para las investigaciones epidemiológicas**

**Piloto's statistics: a digital-educational statistical package for  
the epidemiological researches**

**Manuel Piloto Morejón<sup>1</sup>**

<sup>1</sup>Doctor en Medicina. Especialista de Segundo Grado. Profesor Auxiliar. Metodólogo de Investigaciones. Universidad de Ciencias Médicas de Pinar del Río.

---

## RESUMEN

**Introducción:** El uso práctico de la Bioestadística es una preocupación de cualquier investigador. **Objetivo:** Con el objetivo de facilitar el cálculo matemático de las pruebas estadísticas utilizadas en las investigaciones epidemiológicas, su aplicación e interpretación, se confeccionaron cuatro herramientas, utilizando el software libre y no propietario OpenOffice.org.Calc 2.3 del sistema operativo Unix Ubuntu 8.04 de Linux (con copia a Microsoft Office Excel para su generalización) y con el título: "*ESTADÍSTICA PILOTO. PAQUETE ESTADÍSTICO DIGITAL EDUCACIONAL PARA LAS INVESTIGACIONES EPIDEMIOLÓGICAS*", en Pinar del Río en el 2008. **Material y método:** el software contiene 18 pruebas estadísticas: Ji cuadrado (incluye Corrección de Yates), test de probabilidad exacta de Fisher, Odds Ratio, riesgo relativo, sensibilidad, especificidad, valor predictivo de la prueba positiva y negativa, riesgo atribuible, riesgo atribuible porcentual e índices de Kappa y Youden, pruebas dependientes de los valores a-b-c-d de la tabla de contingencia 2x2, además, el cálculo de media, desviación estándar, varianza, estadígrafo Z y Ji cuadrado de Mc Nemar. **Resultados:** las herramientas fueron validadas mediante programas profesionales y generalizados en trabajos de terminación de las especialidades y maestrías y en otras investigaciones de la provincia. Las herramientas fueron presentadas en los eventos provinciales nacionales y en Sudáfrica (2007) y son las médulas de los cursos de postgrado de Bioestadística aplicada del autor, sin sustituir los programas profesionales, sino que se suman a ellos para facilitar el aprendizaje de la Bioestadística. Sus ventajas son la sencillez, aplicabilidad e inmediatez, los resultados al unísono, la facilidad de interpretación, ayuda en español y su pequeño tamaño.

**Palabras clave:** BIOESTADÍSTICA/instrumentación/métodos, ESTADÍSTICAS VITALES, EDUCACION CONTINUA, APRENDIZAJE, INVESTIGACIÓN BIOMÉDICA.

---

## ABSTRACT

The practical use of Biostatistics is a concern for any researcher. Objective: Four tools were created aimed at providing the mathematical calculation of the statistical tests used in epidemiological researches as well as its application and interpretation, these tools use a free software, without property OpenOffice.org.Calc 2.3 of the Unix Ubuntu 8.04 of Linux operative system (with a copy in Microsoft Office Excel to its generalization) having the title of: "Pilot's statistics. A digital-educational statistical package for the epidemiological researches" in Pinar del Rio during 2008 Material and Method: the software contains 18 statistical tests; chi square test (includes Yates's correction), Fisher's test of exact probability, odds ratio, relative risk, attributable risk percent and Kappa's and Youden's rates, dependent tests a-b-c-d of contingency table 2x2. Besides; the calculation of the mean, standard deviation, variance, Z statistic and Mc Nemar's chi square test. Results: the tools were validated by means of professional programs and generalized in the final researches of the medical specialties, master's degree and other researches of the province. The tools were presented in provincial, national events and in South Africa (2007) and they are fundamental topics in postgraduate courses of Applied Biostatistics to the tutor, without the substitution of the professional syllabuses; the tools are added to easier the acquisition of knowledge in Biostatistics. The most important advantages of these tools are: the simplicity, applicability and immediacy, as well as the results in unison, easy interpretation, help in Spanish language and little size.

**Key words:** BIOSTATISTICS/instrumentation/methods, VITAL STATISTICS, CONTINUING EDUCATION; LEARNING, BIOMEDICAL RESEARCH.

---

## **INTRODUCCIÓN**

La Bioestadística es una ciencia indispensable para el médico en su labor científica, tanto en el trabajo diario, en el momento de emitir un hipotético juicio o diagnóstico positivo o diferencial, como al tratar de planificar, y posteriormente, al concluir una investigación sobre un problema de salud.

Durante años se ha estudiado y trabajado intensamente en el tema, con el objetivo de mejorar los métodos de aplicación-enseñanza-aprendizaje en esa especialidad y en el trabajo diario como profesores universitarios.

Todo médico máxime si de un profesor universitario se trata, debe saber proyectar una investigación en todas sus partes. En ocasiones, se necesita una "asesoría" desde el punto de vista económico, ortográfico, gramatical, bioestadístico, de diseño, u otro tipo, pero todo investigador debe conocer todos los aspectos básicos de la Metodología de la Investigación y la Bioestadística y no bastan los cursos básicos sobre Estadística que se nos exigen para los cambios de categoría docente.

En estudios en Cuba para evaluar el impacto de la especialidad de Bioestadística sobre la esfera intelectual de los egresados se encontró que el 94,7%, el 86%, el 89,5 %, y el 70,2% opinaron que la especialidad influyó mucho en su nivel científico, en su capacidad para realizar proyectos de investigación, comunicar los resultados encontrados e impartir docencia respectivamente; se produjo un incremento en el promedio anual de proyectos de investigación realizados por egresado (0,08 vs 0,32), de asistencia a eventos científicos (0,90 vs 1,40) y de publicaciones realizadas (0,22 vs 1,03). Finalmente, el 76,8% habían iniciado o realizado algún tipo de superación profesional después de cursar la especialidad. Se concluyó que la especialidad tuvo impacto en la esfera intelectual de los egresados.<sup>1,2</sup>

En otro estudio de los mismos autores para evaluar el impacto de la especialidad en Bioestadística sobre su claustro y la Salud Pública Cubana, los expertos opinaron que el impacto de la especialidad en la Salud Pública Cubana se aprecia en la calidad de la información estadística que se brinda en todos los niveles del Sistema Nacional de Salud, en el aumento de la calidad de las investigaciones y eventos científicos realizados en el sector de la salud, en la elevación de la calidad como investigadores y docentes de otros profesionales de la salud, así como en la realización de eventos científicos.<sup>3</sup>

Otro estudio reciente realizado con los residentes de Medicina en los Estados Unidos se concluyó que la mayoría de ellos carecen de conocimientos de Bioestadística para interpretar los resultados de las investigaciones clínicas publicadas y que los programas de estudio debían incluir más entrenamiento efectivo para preparar a los residentes en esta importante habilidad.<sup>4</sup>

Cuando se iniciaron los primeros pasos en la Bioestadística, se utilizó una calculadora portátil y así se enseñaba a los residentes y a otros especialistas jóvenes interesados en el trabajo de la Investigación Biomédica.

Con los conocimientos adquiridos durante años en materia de la Computación, se ha trabajado con el software libre y no propietario OpenOffice.org.Calc 2.3 del sistema operativo Unix, Ubuntu 8.04 de Linux, el cual permite aplicar funciones diversas y así surgió la idea de confeccionar estas herramientas como valiosos instrumentos para facilitar la engorrosa tarea del cálculo de las fórmulas estadísticas de uso frecuente muchas veces complicadas mediante la calculadora, para dar respuesta a una necesidad de los investigadores.

Es innegable la importancia de la tabla de 2x2 en la Bioestadística moderna. El planteamiento de la misma, a partir de tablas más complejas, obedece a la necesidad de obtener 2 filas y columnas, indispensables para aplicar las fórmulas de las pruebas estadísticas de corte epidemiológico para realizar estudios de casos-controles, de cohortes y ensayos clínicos.<sup>5,6</sup>

Cálculos estadísticos, como la prueba de Ji cuadrado (en sus diferentes variantes), el test de la probabilidad exacta de Fischer, el riesgo relativo o el Odds Ratio (para conocer los diferentes factores de riesgo), la sensibilidad o la especificidad (para valorar la eficacia de una prueba diagnóstica), el valor predictivo de la prueba positiva (VPPP) o negativa (VPPN) y el riesgo atribuible (RA) entre otras, se tornan difíciles y trabajosos, máxime si deben acompañarse del cálculo del intervalo de confianza.<sup>7,8</sup>

En el desarrollo de cualquier investigación es importante describir los métodos estadísticos con los detalles suficientes para permitir que un lector versado en el tema y con acceso a los datos originales, pueda verificar los resultados publicados. Además, se debe poner de manifiesto la existencia o no de asociación entre diversas variables. La asociación encontrada puede ser real; sin embargo, con mayor frecuencia de la que uno se imagina ésta es producto del azar, de la existencia de sesgos, de la presencia de variables de confusión o de la variabilidad biológica del fenómeno en estudio. Para dilucidar este problema existen una serie de pasos fundamentales al momento de diseñar y conducir una investigación y, posteriormente, al momento del análisis de los datos es donde aparece la utilización de herramientas estadísticas tanto de carácter descriptivo como analítico.<sup>10,11</sup>

Como objetivo general se trazó confeccionar un paquete de cuatro herramientas digitales en el software libre y no propietario OpenOffice.org.Calc 2.3 en la plataforma Ubuntu 8.04 de Linux, que facilite y agilice el cálculo matemático de las pruebas estadísticas utilizadas en las investigaciones epidemiológicas, dependientes todas de los valores a-b-c-d de la tabla de contingencia de 2x2, la media (en series agrupadas o no), la desviación estándar, la varianza, el estadígrafo Z y el Ji cuadrado de Mc Nemar, validar los resultados de los cálculos estadísticos mediante otros programas profesionales, generalizar estas herramientas mediante su utilización en trabajos de terminación de las especialidades y de maestrías, en otras investigaciones en salud, su presentación en eventos provinciales, nacionales e internacionales y su publicación.

## **MATERIAL Y MÉTODO**

Se confeccionaron cuatro herramientas digitales en el software libre y no propietario (gratis y de libre distribución) OpenOffice.org.Calc 2.3 de la plataforma Ubuntu 8.04 de Linux (sistema operativo multi-área y multi-usuario)

(con copia a Microsoft Office Excel para su generalización) para el cálculo de las 18 pruebas estadísticas.<sup>12,13</sup>

El paquete contiene cuatro herramientas (libros) que son:

"Piloto 2x2 y más" (es el más complejo y contiene varias hojas de cálculo)

"Piloto Z" (contiene 1 hoja de cálculo)

"Piloto Media-DS-Varianza" (contiene 4 hojas de cálculo).

"Piloto Mc Nemar" (contiene 1 hoja de cálculo)

Se elaboró un "Manual de Usuario" y protegieron todas las celdas de cada hoja de cálculo mediante una contraseña, excepto aquellas que permiten al usuario introducir los datos primarios y se colorearon de amarillo para su más rápido reconocimiento.

Los métodos teóricos utilizados fueron:

Revisión bibliográfica. Se realizó una búsqueda bibliográfica para conocer el estado de la temática en la esfera nacional e internacional.

Histórico-lógicos. Posteriormente a la confección de las herramientas, se aplicaron métodos histórico-lógicos en la revisión de documentos referidos al uso de software profesional existentes para conocer su funcionamiento y proceder a la validación de los resultados de las pruebas estadísticas.

Análisis-síntesis para el proceso de diseño e implementación de las herramientas.

Los resultados de las diferentes pruebas fueron validados mediante otros programas digitales profesionales entre los que podemos citar: The TWOBYTWO Analyzer, Microsta, EPI-INFO y SPSS.

En las herramientas se tuvieron en cuenta:

EL ADMINISTRADOR, en este caso el autor, que controla la interfaz diseñada de forma homogénea e intuitiva y protege las celdas mediante una contraseña.

EL USUARIO, que utiliza las herramientas mediante los derechos exclusivos otorgados.

## **RESULTADOS**

Con el título de: Estadística Piloto. Paquete estadístico digital educacional para las investigaciones epidemiológicas: se presentó este trabajo para el cálculo de 18 pruebas estadísticas de uso muy frecuente y 13 de ellas dependientes de los valores de a-b-c-d que resultan de la conocida tabla de 2x2 (2 filas y 2 columnas) en el caso de la herramienta principal.

El paquete se denomina "estadístico" porque realiza el cálculo matemático de las diferentes pruebas; se considera "digital" porque es un software y es "educacional", primeramente se utiliza para calcular, y ofrecer un diagnóstico e interpretar el resultado de las pruebas, pero al mismo tiempo se educa al usuario.

El trabajo presentado es de utilidad para los que dominan el cálculo y la interpretación de las diferentes pruebas estadísticas, pero también para los que tienen poco dominio, porque contiene una ayuda de hipervínculo (en español) de ser necesaria una consulta.

Estas herramientas pueden utilizarse para el cálculo de las siguientes fórmulas:

Ji cuadrado para un grado de libertad (GL) en tablas de 2x2.

Ji cuadrado para 2-16 GL en tablas de >2x2.

Ji cuadrado con corrección de Yates (si  $n < 25$ ) en tabla de 2x2.

Test de probabilidad exacta de Fisher) en tabla de 2x2.

Ji cuadrado (Mc Nemar).

Razón de productos cruzados (Odds Ratio = OR) para estudios epidemiológicos de casos-contrroles (retrospectivos) y su intervalo de confianza al 95%.

Riesgo relativo (RR) (para estudios epidemiológicos de cohorte (prospectivos) y su intervalo de confianza al 95%.

Sensibilidad.

Especificidad.

Valor predictivo de la prueba positiva (VPPP).

Valor predictivo de la prueba negativa (VPPN).

Riesgo atribuible (RA).

Riesgo atribuible porcentual.

Índice de Kappa (para determinar si el grado de concordancia entre el resultado obtenido por un procedimiento diagnóstico y la verdad (presencia o ausencia de patología), excede lo que se esperaría por el azar.

Índice de Youden (para evaluar la eficacia de las pruebas de sensibilidad y especificidad).

Estadígrafo Z.

Cálculo de la media, la desviación estándar (DS) y la varianza.

En la herramienta principal denominada "Piloto 2x2 y más" solamente se le permite al usuario utilizar las cuatro celdas coloreadas de amarillo en las que escribirá los datos numéricos correspondientes a los valores: a-b-c-d. El resto de las celdas donde funciona todo el cálculo estadístico está protegido por una contraseña que no permite dañar las fórmulas (o funciones) del trabajo). Al concluir su utilización y cerrar la herramienta, la computadora siempre preguntará al usuario si desea guardar los cambios y la respuesta afirmativa o negativa utilizada, nunca dañará el contenido del programa.

En las otras tres herramientas existen también celdas coloreadas de amarillo para el uso del usuario (ver Manual del Usuario)

## **DISCUSIÓN**

Las particularidades de esta versión consisten que al introducir los datos a-b-c-d de la tabla de 2x2, de inmediato la computadora calcula el resultado de las pruebas, todas al unísono. El conocedor de la materia sabe que, o se utiliza el Ji cuadrado o el Ji cuadrado con corrección de Yates, o el test de probabilidad exacta de Fisher, pero nunca los 3 juntos; se puede utilizar la sensibilidad y la especificidad al mismo tiempo o se utiliza el Odds Ratio (OR), el riesgo relativo (RR), porque depende si el estudio es retrospectivo o prospectivo respectivamente. De todas formas, todos los resultados aparecen en la pantalla y el usuario utiliza solamente aquel que necesita. Si se cambia alguno de los datos primarios (valores a-b-c-d) la computadora recalculará automáticamente los resultados.

Las herramientas están confeccionadas en una ambiente Web con uso de hipervínculos, para que el usuario pueda navegar (hacia delante o hacia atrás) por el mismo solicitando las opciones que se le ofrecen.

Ventajas del paquete estadístico:

Lo elemental y sencillo de su manipulación mediante el manual del usuario que ayuda a su funcionamiento.

Es un instrumento útil para cualquier investigador, ya que resulta fácil de manejar por estar confeccionado en idioma Español y por contar con una ayuda (también en Español) para conocer el objetivo, el resultado, el diagnóstico y la interpretación de cada prueba. La ayuda de cada prueba estadística (mediante hipervínculo) es inmediata, sólo con un clic en la celda en la que aparece el nombre de la prueba.

La aplicabilidad. El usuario tiene acceso a una interpretación detallada de cada prueba. Por ejemplo: en el caso del Ji cuadrado se obtiene: el valor de Ji, el valor de "p" y el diagnóstico-interpretación del resultado.

La inmediatez de los resultados de las 18 pruebas estadísticas que garantiza la economía del tiempo, elemento de vital importancia para el usuario. Trece pruebas de estadísticas están contenidas en la herramienta principal denominada "Piloto 2x2 y más" (todas dependientes de los valores a-b-c-d de los estudios epidemiológicos), que se calculan al unísono y entonces el usuario decide cuál(es) de ella(s) utilizar y 5 contenidas en las otras 3 herramientas.

Está protegido contra escritura indeseada y/o casual (excepto las celdas coloreadas de amarillo, únicas en las que el usuario puede escribir).

No necesita instalación: puede trabajarse desde un dispositivo externo USB (flash memomry) o desde un CD.

Su pequeño tamaño (Kb).

Validación: el paquete estadístico desde su inicio ha sido validado por otros programas profesionales existentes ofertados por los diferentes colegas, al conocer de la existencia de este trabajo y consultado a otros profesionales de la especialidad de Bioestadística, además, de haber sido utilizado en los últimos 10 años, primeramente para los trabajos de terminación de la especialidad (TTE) de

Ginecología y Obstetricia, y posteriormente, en otras especialidades médicas, en Licenciatura de Enfermería y en las Tesis de Maestrías y Doctorales, y por último, ya fue incorporado al Tema II de la asignatura de Salud Pública (5to año) para la explicación e interpretación de los factores de riesgo y en la asignatura de Bioestadística de los primeros años de las diferentes carreras.

Ha sido aplicado en Cuba y en Sudáfrica donde fue presentado como conferencia magistral (actividad plenaria de clausura del Congreso de las Maestrías en Sudáfrica-2007).

Se consideraron como impacto los siguientes:

Social. Esta tecnología tiene impacto en los investigadores, ya que su aprendizaje es posible y su generalización no es difícil (de hecho, ya se ha estado generalizando desde hace años).

Tecnológico. Son herramientas de fácil acceso a principiantes o conocedores y su aplicación se inserta dentro de las políticas, programas y objetivos de salud establecidos por el país, ya que es efectiva y segura.

Científico. Es aplicable a cualquier nivel de conocimiento de Estadística, desde alumnos de 1ro, 2do y 5to año de las diferentes carreras de Medicina hasta el de técnicos y profesionales de la Salud.

Económico. Produce una disminución en el empleo de tiempo y recursos de los usuarios con implicaciones positivas con respecto a la productividad del trabajo en sus investigaciones.

Se confeccionaron cuatro herramientas digitales (software) para el cálculo matemático inmediato y al unísono de la mayoría de las diferentes pruebas estadísticas utilizadas en las investigaciones epidemiológicas, aprovechando la aplicación fácil e interpretación para su aprendizaje mediante un Manual del Usuario y una ayuda en español.

Se puede afirmar, que el software confeccionado reúne cualidades como: la efectividad, pertinencia, utilidad, seguridad, eficiencia, el impacto socio-económico e implicación organizativa.

Este software está protegido por los Derechos de Autor: © Copyright Manuel Piloto Morejón 2009 y obtuvo entre otros galardones: el Premio Anual de la Salud 2009 y el Premio de la Academia de Ciencias 2009.

¿Cómo acceder al software?

El usuario puede acceder al software (comprimido) mediante un clic encima del link que a continuación se oferta: [software Estadística Piloto.rar](#)

Una vez descargada esta carpeta comprimida a un lugar por designado, deberá ser descompactada y de inmediato el software estará a su disposición.



## REFERENCIAS BIBLIOGRÁFICAS

1. Couturejuzón González Lourdes, Columbié Pileta Miday. Impacto en la esfera intelectual del egresado de la especialidad de Bioestadística. Educ Med Super [revista en Internet]. 2008 Dic [citado 2010 Jul 06]; 22(4): [aprox. 7 pantallas]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-21412008000400004&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412008000400004&lng=es).
2. Couturejuzón González Lourdes, Columbié Pileta Miday. Impacto en la esfera intelectual del egresado de la especialidad de Bioestadística. Educ Med Super [revista en Internet]. 2009 Mar [citado 2010 Jul 07]; 23(1): [aprox. 6 pantallas]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-21412009000100003&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412009000100003&lng=es).
3. Couturejuzón González Lourdes, Columbié Pileta Miday. Impacto de la Especialidad en Bioestadística sobre su claustro y sobre la Salud Pública Cubana, 1974-2006. Rev Cubana Salud Pública [revista en Internet]. 2009 Mar [citado 2010 Jul 06]; 35(1). [aprox. 8 pantallas]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-34662009000100012&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662009000100012&lng=es).
4. Windish DM; Huot SJ; Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA [revista en Internet] 2007 Sep 5 [citado 2010 Jul 08]; 298(9):1010-22. Disponible en: <http://jama.ama-assn.org/cgi/content/full/298/9/1010?maxtoshow=&hits=10&RESULTFORMAT=&fulltext=Medicine+residents+understandings&searchid=1&FIRSTINDEX=0&resourcetype=HWCIT>
5. Boulesteix AL. Maximally selected chi-square statistics and binary splits of nominal variables. Biom J; [revista en Internet] 2006 Aug [citado 2010 Jul 08]; 48(5):838-48. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/17094347>
6. Boulesteix AL. Maximally selected chi-square statistics for ordinal variables. Biom J; [revista en Internet] 2006 Jun [citado 2010 Jul 08] 48(3):451-62. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/16845908>
7. Colectivo de autores. Metodología de la Investigación Educativa. Desafíos y polémicas actuales. Habana: Ed Ciencias Médicas, 2003. En: <http://www.universitas.net.ve/biblioteca/ELPROCESODEINVESTIGACION.doc>
8. Fathalla MF, Fathalla MMF. Descripción y análisis de los resultados de investigación. Interpretación de los resultados de investigación. En: Guía práctica de investigación en Salud. Publicación científica y técnica No.620. OPS; Washington DC; E.U.A. 2008.p.86-108.
9. Díaz Novás José, Gallego Machado Bárbara. Algunas medidas de utilidad en el diagnóstico. Rev Cubana Med Gen Integr [revista en Internet]. 2006 Mar [citado 2010 Jul 07]; 22(1). [aprox. 8 pantallas]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0864-21252006000100008&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21252006000100008&lng=es).
10. Artilles VL, Otero Iglesias J, Barrios Osuna I. Estadística. En: Metodología de la Investigación para las Ciencias de la Salud. Edit. Ciencias Médicas. La Habana; 2009.p.312-13.

11. Manterota DC, Pineda NV. El valor de "p" y la "significación estadística". Aspectos generales y su valor en la práctica clínica. Rev Chil Cir [revista en Internet]. 2008 feb. [citado 2010 Jul 08]60(1). [aprox. 8 pantallas].

Recibido: 12 de julio de 2010.

Aprobado: 7 de septiembre de 2010.

Dr. Manuel Piloto Morejón. Calle Rafael Morales #86 (Norte). Pinar del Río.

Teléfono: 75-3108.

E- mail: [piloto@princesa.pri.sld.cu](mailto:piloto@princesa.pri.sld.cu)