
Software de Aplicación Portable como herramienta comprobatoria de las medidas de dispersión para datos no agrupados en la Estadística Descriptiva.

Espinoza Maza, Jonathan de Jesús
Universidad Veracruzana, México
jespinoza@uv.mx

Fragoso Terán, Juan Manuel
Universidad Veracruzana, México
juafragoso@uv.mx

Rodríguez Rodríguez, Luis Alberto
Universidad Veracruzana, México
alberodriguez@uv.mx

Aguirre Orozco, Ulises Gonzalo
Universidad Veracruzana, México
uaguirre@uv.mx

Mendoza Barrón, Gustavo
Universidad Veracruzana, México
gusmendoza@uv.mx

Resumen — Las medidas de dispersión indican la variación de los datos de una distribución, mostrando por medio de un valor, que tan alejados están los datos de un espacio muestral. Conocer las medidas de dispersión permite saber si los datos son parecidos o tienen variaciones significativas. Este artículo, pretende aportar al conocimiento un Software de Aplicación Portable desarrollado con Lenguaje C/C++, a fin de que contribuya con la obtención de resultados para datos no agrupados a través de la Ciencia Computacional.

Palabras claves — Medidas de dispersión; rango; varianza; desviación estándar; coeficiente; variación; media; valores; software;

Abstract — Dispersion measures indicate the variability of the data in a distribution, showing by means of a value, how spread are the data of a sample space. The understanding of dispersion measures allows us to know if the data are similar or have significant variations. This article intends to contribute to the realm of knowledge a portable application developed using the C/C++ programming language, in order to contribute with obtaining results for non-grouped data through Computational Science.

Keywords — Measures of dispersion; rank; variance standard deviation; coefficient; variation; half; values; software;

INTRODUCCIÓN

Hoy en día la pertinencia de utilizar Software de aplicación en las aulas es cada día más importante y necesaria para complementar la formación de los educandos, es por eso que día a día se deben explorar y crear herramientas que sean de utilidad para poder realizar comprobaciones, para este caso, funciones Estadísticas, mismas que se centran en medidas de dispersión para datos no agrupados.

OBJETIVOS

General

Crear un software de aplicación portable que sirva como herramienta comprobatoria para el uso en Estudiantes o usuarios de la Estadística.

Específicos

1. Actualización constante para mejorar presentación y funcionalidad del Software de Aplicación.
2. Contribuir con esta aplicación con el fin de aportar una parte modular de un Software Estadístico y Probabilístico más robusto.

METODOLOGÍA DE DESARROLLO.

Para la elaboración de este Software se aplican los principios de la Metodología Ágil con Programación Extrema.

En el caso de la Metodología XP (extreme programming), permite proveer el software que el cliente necesita y cuando lo necesita, por lo tanto, esta metodología trata de responder de manera ligera, eficiente, con bajo riesgo, flexible, predecible y cómoda para el desarrollo del software, además de que permite potenciar al máximo el trabajo colaborativo; desde el equipo desarrollador hasta el cliente con una retroalimentación efectiva. En este caso la figura del cliente es esencialmente todo aquel estudiante de la Estadística descriptiva, respecto al tema de las medidas de dispersión para datos no agrupados, lo cual permite hacer una retroalimentación ágil y eficaz respecto al binomio estudiante-desarrollador. (Delgado Expósito, 2008)

XP tiene los siguientes principios genéricos (Ingeniería del Software y Sistemas de Información, 2003):

- I. La prioridad es satisfacer al cliente mediante tempranas y continuas entregas de software que le aporte un valor.
- II. Dar la bienvenida a los cambios. Se capturan los cambios para que el cliente tenga una ventaja competitiva.
- III. Entregar frecuentemente software que funcione desde un par de semanas a un par de meses, con el menor intervalo de tiempo posible entre entregas.
- IV. La gente del negocio y los desarrolladores deben trabajar juntos a lo largo del proyecto.
- V. Construir el proyecto en torno a individuos motivados. Darles el entorno y el apoyo que necesitan y confiar en ellos para conseguir finalizar el trabajo.
- VI. El diálogo cara a cara es el método más eficiente y efectivo para comunicar información dentro de un equipo de desarrollo.
- VII. El software que funciona es la medida principal de progreso.
- VIII. Los procesos ágiles promueven un desarrollo sosteni-

ble. Los promotores, desarrolladores y usuarios deberían ser capaces de mantener una paz constante.

- IX. La atención continua a la calidad técnica y al buen diseño mejora la agilidad.
- X. La simplicidad es esencial.
- XI. Las mejores arquitecturas, requisitos y diseños surgen de los equipos organizados por sí mismos.
- XII. En intervalos regulares, el equipo reflexiona respecto a cómo llegar a ser más efectivo, y según esto ajusta su comportamiento.

La ventaja de la metodología ágil, es que es fácilmente migrable a una metodología más robusta, tal como el ciclo de vida de desarrollo de software utilizando MDD¹.

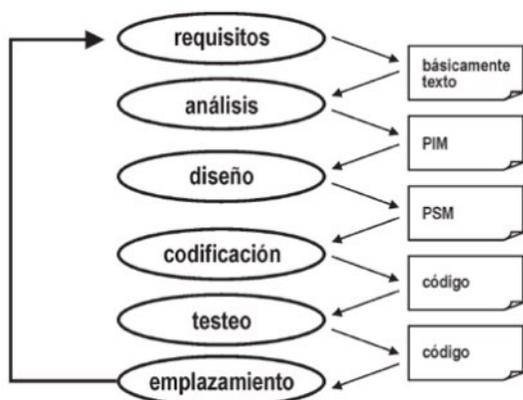
La diferencia de utilizar el ciclo de vida tradicional con el ciclo de vida MDD radica en que en este último se crean modelos formales, es decir, modelos que pueden ser comprendidos por una computadora (más allá de la documentación y pseudocódigo).

¹ Model-Driven Software Development (MDD), que en español se traduce como desarrollo de software dirigido por modelos. (Pons, Giandini, & Pérez, 2010)

MDD identifica distintos tipos de modelos que se pueden aplicar en la fabricación de un Software más robusto y más completo en cuanto a la Estadística y Probabilidad:

- Modelos con alto nivel de abstracción independientes de cualquier metodología computacional, llamados CIMs (Computational Independent Model),
- Modelos independientes de cualquier tecnología de implementación llamados PIMs (Platform Independent Model)
- Modelos que especifican el sistema en términos de construcciones de implementación disponibles en alguna tecnología específica, conocidos como PSMs (Platform Specific Model),

Ilustración 1. Ciclo de vida del desarrollo de software dirigido por modelos.



Fuente: Pons, Giandini, & Pérez, 2010.

- y finalmente modelos que representan el código fuente en sí mismo, identificados como IMs (Implementation Model). (Pons, Giandini, & Pérez, 2010)

Sin lugar a dudas la Metodología de Desarrollo de Software es muy variada y depende en gran medida de las circunstancias en las que se presenten los proyectos, para este fin académico, se requiere agilidad y flexibilidad, con la finalidad de robustecer el desarrollo con el paso del tiempo.

MEDIDAS DE DISPERSIÓN.

La primera medida a analizar es el “rango”, la cual es la medida más simple de variabilidad o dispersión y es tan sencilla como obtener la diferencia entre el valor muestral más grande y el más pequeño. Una fórmula sencilla por la que se puede representar es:

$$R = \text{Max } x_i - \text{Min } x_i$$

Ecuación 1. Fórmula para obtener el Rango.

Donde:

R =Rango

$\text{Max } x_i$ = Valor muestral más grande

Min x_i = Valor muestral más pequeño

La siguiente medida de dispersión se trata de la “**varianza**”, la cual posee como principal característica el poder cuantificar la variación entre los datos de un espacio muestral de forma cuadrática que con otro tipo de medida (como la **media**) no se podría analizar, y es que detectar la variabilidad en un conjunto de datos con solo tomar como referencia la media es una cuestión compleja, por lo que se deberá analizar la varianza. Se pretende que la medición tenga valor positivo alto, si la variable aleatoria fluctúa en el sentido de que frecuentemente asume valores que distan mucho de la media; debe tener un valor positivo pequeño si los valores de X tienden a agruparse muy cerca de la media. Son varias las maneras de definir esa medición. (Milton & Arnold, 2004).

Para efectos de identificación, denotaremos la varianza con la función s^2 .

Para calcular la varianza para datos no agrupados es necesario apoyarse con la siguiente fórmula:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

Ecuación 2. Fórmula para obtener la varianza

Donde:

s^2 = Varianza o Varianza muestral

\sum = Sigma o Sumatoria

x_i = Cada uno de los datos de la distribución

\bar{x} = Media Aritmética

n = Total de los datos de la distribución

La tercera medida es la “**desviación estándar**”, misma que analiza la población y su variación en los datos de la muestra, permitiendo convertir a unidades lineales la dispersión de los datos, teniendo así, las mismas unidades de medición del resto del conjunto de datos, es decir, convierte la varianza, la cual está expresada en unidades cuadradas a unidades lineales.

Esta se da mediante la siguiente fórmula:

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}}$$

Ecuación 3. Fórmula para obtener la Desviación Estándar (Forma Completa)

O de manera más abstracta y sencilla de entender:

$$s = \sqrt{s^2}$$

Ecuación 4. Fórmula para obtener la Desviación Estándar (Forma Corta).

Donde:

s = Desviación Estándar

s^2 = Varianza

Es importante mencionar que el hecho de emplear a $n-1$ en la expresión para determinar la varianza en lugar de n en el denominador de las expresiones anteriores, se debe a que el uso de n tiende a subestimar la varianza y la desviación estándar poblacionales. Empleando $n-1$, se obtiene una corrección adecuada de esa tendencia o sesgo.

Por último, se analiza el “**Coficiente de Variación o Coficiente de Variación de Pearson**”, el cual analiza el problema de las medidas de dispersión al momento de desear comparar las muestras de variables entre sí y no tienen las mismas unidades, justo en ese momento se recurre a las medidas de dispersión relativas, las cuales permiten comparar dos o más distribuciones que no poseen las mismas

unidades o en su defecto, cuando las medias no son iguales. En otras palabras, estamos hablando de la desviación estándar expresada como porcentaje de la media aritmética, y la cual expresa también la relación entre la media y la desviación estándar.

El coeficiente de Variación de Pearson se formula de la siguiente manera.

$$r = \frac{S}{\bar{x}}$$

Ecuación 5. Fórmula para obtener el Coeficiente de Variación de Pearson.

Donde:

r = Coeficiente de Variación de Pearson o Coeficiente de Correlación Lineal de Pearson

s = Desviación Estándar

\bar{x} = Media del Conjunto de Observaciones

EL SOFTWARE COMO HERRAMIENTA DE APOYO AL APRENDIZAJE.

Existen diversas formas de explicar el Software educativo, sin embargo, todas coinciden en que es una herramienta de apoyo al proceso de enseñanza – aprendizaje. Citando una definición del autor Ceja ME-

NA, dice que el Software Educativo “son aquellos programas creados con la finalidad específica de ser utilizados como medio didáctico; es decir, para facilitar los procesos de enseñanza y de aprendizaje tanto en su modalidad tradicional presencial, como en la flexible y a distancia” (Fernández Aedo & Delavaut Romero, 2008).

Por lo que se considera necesario continuar reforzando la creación y utilización de Software de Aplicación orientados a la Estadística, y aunque hay un sin número de ellos, existen diversos inconvenientes, que van desde el grado de comple-

jidad al utilizarlos, hasta el licenciamiento, puesto que muchas de estas aplicaciones son hechas con fines lucrativos.

Para efectos didácticos y con el fin de tener una herramienta comprobatoria para obtener las medidas de variabilidad de datos no agrupados, se ha desarrollado un Software comprobatorio en Lenguaje C/C++ con la finalidad de contar con una herramienta que mejore el binomio enseñanza – aprendizaje y que pueda ser mejorado con el tiempo, incluso migrado a una Aplicación Móvil o con otro Lenguaje de desarrollo como Python, por ejemplo.

Ilustración 2. Entrada del número de datos a procesar.

```

=====
-----Medidas de Dispersión-----
Ingrese la cantidad de datos a procesar
8

```

Fuente: Elaboración propia.

Esta aplicación portable está diseñada sobre Lenguaje C/C++, el cual ofrece las mismas características inherentes propias de desarrollar en dicho lenguaje, así como sus ventajas, como ser un lenguaje de propósito general, conciso, rico en operadores, recursivo, compilación rápida y compacta, portable, y potente. (Badenas Carpio, Llopis Borrás, & Coltell Simón, 2001)

La aplicación comprobatoria desarrollada es sencilla de operar, la cual consta de los pasos mostrados en las ilustraciones 2 a 4.

La aplicación hace el resto. A través de diversas operaciones, ordena los datos de Mayor a Menor y viceversa, arrojando el Rango, la Media, la Varianza, la Desviación Estándar y el Coeficiente de Variación de Pearson.

De esta manera se obtienen las medidas de variabilidad que se desean conocer con ayuda del Software de Aplicación desarrollado.

Para obtener este Software de Aplicación, se debe ingresar al sitio personal del primer autor: <http://jespinoza.byethost8.com/wp/> e ingresar a la sección “Aplicaciones”.

Ilustración 3. Entrada de cada uno de los datos.

```

=====
-----Medidas de Dispersión-----
Ingrese la cantidad de datos a procesar
8
Ingrese el dato 1 30
Ingrese el dato 2 20
Ingrese el dato 3 15
Ingrese el dato 4 44
Ingrese el dato 5 2
Ingrese el dato 6 47
Ingrese el dato 7 50
Ingrese el dato 8 21

```

Fuente: Elaboración propia.

Ilustración 4. Aplicación desarrollada.

```

Ingrese el dato 5 2
Ingrese el dato 6 47
Ingrese el dato 7 50
Ingrese el dato 8 21
ORDENAMIENTO
Mayor a Menor:
50 , 47 , 44 , 30 , 21 , 20 , 15 , 2 ,

De Menor a Mayor :
2 , 15 , 20 , 21 , 30 , 44 , 47 , 50 ,

El rango es 48
La media es 28.625000
La Varianza es 294.267853
La Desviación Estandar es 17.154238
El Coeficiente de Variación de Pearson es 0.599275
=====
© Mtro. Jonathan de Jesús Espinoza Maza
=====

```

Fuente: Elaboración propia.

Ilustración 5. Instrumento difundido a través de Google Forms.

The screenshot shows a Google Form interface with a purple header. The title of the form is "Evaluación para la Aplicación Portable de Medidas de Variabilidad para Datos no Agrupados". Below the title, there is a subtitle: "Implementación de la aplicación para comprobación de Medidas de Dispersion." The form contains five text input fields, each with a red asterisk indicating it is required:

- Nombre *
- Matrícula *
- Programa Educativo *
- Facultad *
- E-mail *

Each field has a small "Text de respuesta corto" label below it. The form is displayed in a preview mode, with a "PREGUNTAS" / "RESPUESTAS" toggle at the top and a "ENVIAR" button in the top right corner.

Fuente: Gráfico extraído de Google Forms.

RESULTADOS PREVIOS A LA DISTRIBUCIÓN.

Durante la elaboración de este artículo, se realizaron pruebas piloto con el Software para determinar su funcionalidad, utilizando como grupo focal una sección de la Experiencia Educativa (asignatura o materia) Probabilidad y Estadística de la Facultad de Ingeniería Mecánica y Ciencias Navales de la Universidad Veracruzana en la Región Vera-

cruz, esto con la finalidad de obtener una retroalimentación formal de los Estudiantes y de su percepción al utilizar el Software de Aplicación Portable en el tema referente a las medidas de dispersión para datos no agrupados en la Unidad de Estadística, obteniendo los siguientes resultados a través de un instrumento de tipo encuesta, difundido a través de Google Forms.

Ilustración 6. Pregunta relacionada con la simplicidad de la Aplicación.

¿Su uso es simple?

15 respuestas

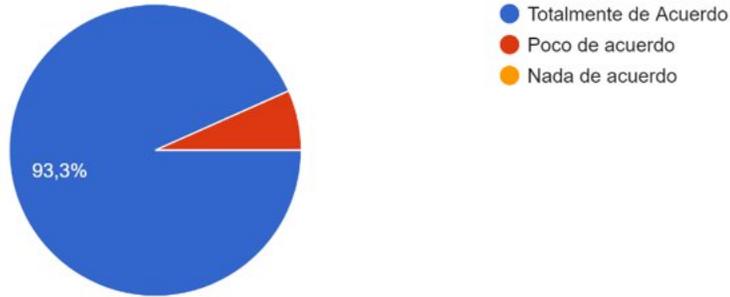


Fuente: Gráfico extraído de Google Forms.

Ilustración 7. Pregunta relacionada con la interfaz de la aplicación.

¿La interfaz es amigable/intuitiva?

15 respuestas



Fuente: Gráfico extraído de Google Forms.

Ilustración 8. Pregunta relacionada con la aportación al conocimiento respecto al tema abordado.

¿La aportación al tema que presenta el sistema es útil?

15 respuestas



Fuente: Gráfico extraído de Google Forms.

Ilustración 9. Pregunta relacionada con la presentación de la información obtenida en la aplicación.

¿La información se presenta de manera clara y entendible?

15 respuestas



Fuente: Gráfico extraído de Google Forms.

Ilustración 10. Pregunta relacionada con la portabilidad de la aplicación y la no dependencia del internet para su funcionamiento.

¿Está de acuerdo con que esta aplicación sea portable? (Es decir, que la pueda ejecutar en cualquier computador... requerir internet para su ejecución)

15 respuestas



Fuente: Gráfico extraído de Google Forms.

Ilustración 11. Pregunta relacionada la necesidad de utilizar esta aplicación vía web además de ser portable.

¿Le gustaría que esta aplicación también esté disponible en línea o en su defecto como Aplicación Móvil?

15 respuestas



Fuente: Gráfico extraído de Google Forms.

Tabla 1. Datos cualitativos obtenidos del instrumento aplicado.

Describa que otra aplicación le gustaría que se desarrollara como herramienta comprobatoria para el tema que usted desee.	¿Qué ventajas encuentra utilizar una aplicación como herramienta comprobatoria?	Por favor escriba una opinión constructiva con la finalidad de mejorar continuamente y para efectos de retroalimentación
Desarrollo de análisis dimensional en fluidos	Tener un medio por el cual saber si lo que hacemos esta correcto	La herramienta está completa en lo que se pretende buscar
Aplicación para la comprobación de esfuerzos en Mecánica de Materiales	Para verificar que el resultado obtenido sea correcto.	La aplicación es muy buena. Su uso es fácil e intuitivo.
Alguna aplicación para algoritmos computacionales	Desaparecer la duda de que el resultado sea correcto.	A mi parecer es muy fácil de usar y demasiado útil.
Para problemas de geometría analítica	Es simple	Me parece una buena herramienta
Sin comentarios	Es rápido y muy fácil de utilizar	Sin comentarios
Obtención de la frecuencia	Fácil de utilizar y exacta	Esta bien la aplicación
Un software para las frecuencias	Disminuye el factor de error por un mal cálculo	Una opción para borrar algún dato mal declarado
iPads	Un uso más rápido	Muy buena aplicación
Comprobación de dominio y rango de funcioness de varias variables hay muy pocas	Puedo saber si estoy bien y si son muchos datos me es más fácil verlo en la aplicación	Ningún comentario
Para calcular y testada	Ahorro de tiempo	Que se puedan corregir los datos, sin empezar todo el procedimiento
Tablets,	El uso es más rápido y fácil	Excelente aplicación
Herramientas de programación	Comprobacion	Uso movil
Una aplicación que desarrolló los cálculos utilizados en mecánica de materiales.	Verificar los datos obtenidos mediante el medio convencional	Pienso que, sería muy útil una opción para modificar los datos de entrada sin necesidad de empezar el un nuevo proceso.
Ninguna	Más rapidez a la hora de sacar medidas de dispersión en da datos no agrupados	Colores más sencillos o con buenos contrastes, fuentes de letra más agradables a la vista

Fuente: Elaboración propia.

CONCLUSIÓN

Las medidas de dispersión permiten saber que tan alejados se encuentran los datos de una muestra con respecto a su media aritmética, y entre sí, con la finalidad de obtener un panorama de la conformación de dicho espacio muestral y poder realizar comparaciones o análisis a fondo. Esto con ayuda de todas las demás medidas de dispersión y de la uniformidad de las unidades, siendo un conjunto de herramientas complementarias al análisis estadístico, y que incluso sientan la base de la Estadística Multivariante para obtener Hipótesis Estadísticamente Viables o No viables. El Software de Aplicación orientado a la educación, es un camino y una herramienta más que los profesores y educandos tienen a su alcance como recurso para comprender de mejor manera y aplicar el conocimiento adquirido, además de ser una herramienta comprobatoria con respecto a los resultados obtenidos en las diversas operaciones realizadas.

Los estudiantes hoy en día deben poseer soluciones al alcance, que utilicen desde los recursos de desarrollo más sencillos y ágiles

hasta los que sean orientados a la web, con la finalidad de cerrar la brecha digital y seguir trabajando para mejorar la calidad de la educación y dotarlos de un número mayor de soluciones tecnológicas que les permita incrementar su nivel educativo.

REFERENCIAS

- Badenas Carpio, J., Llopis Borrás, J., & Coltell Simón, Ó. (2001). *Curso Práctico de Programación en C y C++*. Castelló de la Plana: Universitat Jaume I.
- Cairó, O. (2006). *Metodología de la Programación*. Ciudad de México: Alfaomega.
- Deitel, P. J., & Deitek, H. M. (2008). *C++ Cómo Programar*. Ciudad de México: Pearson Prentice Hall.
- Devorce, J. L. (2016). *Probabilidad y Estadística para Ingeniería y Ciencias*. Ciudad de México: Cengage Learning.
- Fernández Aedo, C. R., & Delavaut Romero, M. E. (2008). *Educación y Tecnología. Un binomio excepcional*. Buenos Aires: Grupo Editor K.
- Jiménez Murillo, J. A., Jiménez Hernández, E. M., & Alvarado Zamora, L. N. (2015). *Fundamentos de Programación. Diagramas de flujo, diagramas N-S. Pseudocódigo y Java*. Ciudad de México: Alfaomega.
- Milton, J. S., & Arnold, J. C. (2004). *Probabilidad y Estadística. Con aplicaciones para ingeniería y Ciencias Computacionales*. Ciudad de México: Mc Graw Hill.

Nieves, A., & Domínguez, F. C. (2010).
Probabilidad y Estadística para
Ingeniería. Un Enfoque Moderno. Ciudad
de México: Mc Graw Hill.