

## APRENDIZAJE UBÍCUO DE ESTADÍSTICA DESCRIPTIVA E INFERENCIAL

\* FRANCISCO JAVIER TAPIA MORENO, HÉCTOR ANTONIO VILLA MARTÍNEZ

### RESUMEN

Hemos desarrollado herramientas de enseñanza y aprendizaje para la estadística descriptiva e inferencial para dispositivos móviles que utilizan el sistema operativo Android. Con estas herramientas los estudiantes pueden desarrollar diversas actividades: i) construir y analizar un histograma, un polígono de frecuencias o una gráfica de barras o de sectores, según el tipo de variable aleatoria que se esté procesando; ii) calcular medidas sumarias tales como valores máximos y mínimos, cuartiles, deciles, rango intercuartílico, media aritmética, mediana, moda, desviación estándar y varianza; iii) obtener los coeficientes de regresión y determinación, la ecuación de regresión lineal y la gráfica correspondiente; iv) calcular los coeficientes de asimetría y el coeficiente de curtosis; v) calcular intervalos de confianza para la media aritmética y para la proporción de la población, basada en una muestra grande o pequeña de datos de una población normal, cuasi-normal, o de una no normal. Este trabajo muestra los logros del proyecto "Statistics-to-Go" que estamos desarrollando en el Departamento de Matemáticas de la Universidad de Sonora. Trabajos futuros incluyen el diseño de herramientas para calcular intervalos de confianza para diferentes tareas: la diferencia entre dos medias, diferencia entre dos proporciones, la varianza, el cociente de dos varianzas y para realizar pruebas de hipótesis.

**Palabras clave:** Android, estadística, aprendizaje móvil, aprendizaje ubicuo.

### ABSTRACT

We have developed tools for teaching and learning descriptive and inferential statistics on mobile devices using the Android operating system. With these tools, students can: i) Construct and analyze a histogram, a frequencies polygon, a bar chart, or a pie chart, depending on the type of random variable being processed, ii) computing summary measures such as maximum and minimum values, quartiles, deciles, interquartile range, arithmetic mean, median, mode, standard deviation, and variance, iii) obtain the coefficients of regression and determination, the linear regression equation and the corresponding graph, iv) calculating the coefficients of skewness and kurtosis, v) calculate confidence intervals for the arithmetic mean and the proportion of the population based either on a large or a small data sample from a quasi-normal, non-normal, or normal population. This work shows the achievements of the "Statistics-to-Go" project that we are developing in the Department of Mathematics at the University of Sonora. Future work includes the design of tools to calculate confidence intervals for the difference between two means, difference between two proportions, variance, the ratio of two variances, and to test hypotheses

**Keywords:** Android, statistics, mobile learning, ubiquitous learning.

DR. FRANCISCO JAVIER TAPIA MORENO  
Correo: ftapia@gauss.mat.uson.mx  
M.C. HÉCTOR ANTONIO VILLA MARTÍNEZ  
Correo: hvilla@gauss.mat.uson.mx  
Departamento de Matemáticas.  
Universidad de Sonora

## INTRODUCCIÓN

En la Universidad de Sonora todos los estudiantes de licenciatura toman un curso obligatorio de introducción a la estadística. El objetivo del curso es dotar a los estudiantes de las herramientas estadísticas básicas y que se familiaricen con el análisis estadístico utilizando software estadístico. Estas herramientas de software sólo se ejecutan en computadoras de escritorio o en laptops, lo que significa que los estudiantes deben estar ya sea en el laboratorio de computación o llevar sus computadoras portátiles. Además, la mayoría de estos estudiantes no tienen una computadora, laptop o internet en sus hogares y tienen dificultades para hacer sus tareas y prácticas. Sin embargo, casi todos los estudiantes tienen un teléfono celular, ya que es mucho más barato que una computadora personal o una laptop. Con el propósito de dar una solución a esta grave situación, en septiembre de 2012 sometimos un proyecto denominado "Statistics-to-Go" al Departamento de Matemáticas de la Universidad de Sonora, con el propósito de producir las herramientas para llevar a cabo los cálculos estadísticos de estadística inferencial en teléfonos celulares o tabletas equipados con el sistema operativo Android.

En la primera etapa de este proyecto [1], hemos diseñado herramientas para: i) calcular todas las medidas centrales (media aritmética, mediana, moda y percentiles), dispersión (varianza, desviación estándar y el coeficiente de variación), medidas de forma (coeficientes de asimetría y curtosis); ii) dibujar diagramas circulares, diagramas de barras, histogramas, polígonos de frecuencias, polígonos acumulativos

y diagramas de dispersión; y iii) calcular la ecuación del modelo de regresión lineal y determinar los coeficientes de correlación.

En el resto de este trabajo se presenta la justificación y motivación de nuestra investigación, la definición de aprendizaje móvil y sus principales ventajas sobre otros tipos de aprendizaje, los detalles sobre el sistema operativo Android, la definición de los intervalos de confianza para la media aritmética de la población y algunas formas de calcularlo, los resultados del proyecto "Statistics-to-Go" en su segunda etapa", y por último, se presentan las conclusiones y los trabajos futuros.

## APRENDIZAJE MÓVIL

El aprendizaje móvil (también conocido como *m-learning*) es la adquisición de conocimientos a través de la tecnología de computación móvil, tales como teléfonos celulares y tabletas. Esta modalidad permite a las personas aprender en cualquier momento y en cualquier lugar. Esta última noción "en cualquier momento y en cualquier lugar", aparece con frecuencia descrita como "ubicua" en la literatura sobre Tecnologías de la Información y Comunicación (TIC's), surgiendo así el concepto de aprendizaje ubicuo [2]. Pensamos que en un futuro próximo, este tipo de aprendizaje se utilizará a diario por personas autodidactas.

Las diferencias entre el *m-learning* y otros tipos de aprendizaje pueden percibirse desde dos perspectivas: la tecnológica y la experiencia educativa. En cuanto a la tecnología, el aprendizaje móvil se distingue por el uso de equipo portátil que permite que el estudiante tenga acceso



a herramientas de aprendizaje en cualquier momento y en cualquier lugar. Con respecto a la experiencia educativa, John Traxler [3] define el aprendizaje móvil mediante palabras clave. De esta manera, el aprendizaje móvil es personal, espontáneo, oportunista, informal, omnipresente, privado, contextual, segmentado, y portátil. Traxler también señala que algunas de estas características pueden desaparecer con el progreso de la tecnología móvil. Sin embargo, propiedades tales como la informalidad, la movilidad y el contexto se mantendrán.

Desde el punto de vista de los estudiantes, el aprendizaje móvil posee varias ventajas: 1) La eliminación de las limitaciones de tiempo y espacio. Esto significa que el aprendizaje está disponible en cualquier momento y casi en cualquier lugar. Por lo tanto, el tiempo ocioso se puede utilizar para el aprendizaje; 2) permite la comunicación con sus compañeros y con el profesor sin contacto físico; 3) admite la recepción de objetos de aprendizaje que dependen de la ubicación (es decir, aprendizaje contextual); 4) recibe retroalimentación instantánea a los ejercicios y exámenes; y 5) reconoce la recepción de objetos de aprendizaje que empleen imágenes, video y audio.

## ANDROID

Android es un sistema operativo diseñado principalmente para dispositivos móviles tales como teléfonos celulares y tabletas.

Desarrollado inicialmente por *Android Inc.*, adquirida por *Google* en 2005, Android se dio a conocer en 2007 junto con la creación de la *Open Handset Alliance*, un consorcio de hardware, software y las empresas de telecomunicaciones dedicadas a la promoción de estándares abiertos para dispositivos móviles. El primer teléfono con Android fue vendido en octubre de 2008 [4]. Los teléfonos celulares basados en Android lideran el mercado. De acuerdo con *Tech-thoughts* [5], Android tiene una participación de casi el 70% en el mercado mundial de smartphones, mientras que el *iPhone Operating System* (iOS), fabricado por *Apple*, representa alrededor del 14%. De hecho, Android tiene la mayor participación en todos los países desarrollados, excepto en Estados Unidos y Japón, donde iOS es el sistema operativo móvil dominante [6, 7]. En el mundo de las tabletas la historia es similar, de acuerdo con BGR [8] la cuota de mercado mundial de Android creció de 39.8% en 2011 a 42.7% en 2012, mientras que la participación de iOS disminuyó de 56.3% a 53.8%.

Android tiene muchas ventajas sobre otros sistemas operativos. Aquí mencionamos las más importantes y más congruentes con la necesidad de los estudiantes de

bajos recursos económicos. Según orden de importancia: **i)** es una plataforma de código abierto e integrada. Las empresas pueden publicar actualizaciones y revisiones que los usuarios pueden descargar por su cuenta; **ii)** el kit de desarrollo de software (SDK) de Android permite a cualquiera crear sus propias aplicaciones, debido a que el SDK incluye un conjunto completo de herramientas de desarrollo como depurador, bibliotecas, emulador de terminal, código ejemplo, documentación y tutoriales; **iii)** Android utiliza el lenguaje de programación Java, el cual es un lenguaje estandarizado y abierto. El emulador de la plataforma Android tiene un diseño moderno y es fácil de usar; **iv)** es posible la instalación de todo el entorno para desarrollar aplicaciones de Android en los principales sistemas operativos de escritorio; y **v)** Android permite a los desarrolladores diseñar aplicaciones para los usuarios en función de sus preferencias. Esto proporciona a los usuarios la capacidad de tener experiencias únicas diseñadas específicamente para ellos, ya que Android es altamente personalizable.



## INTERVALOS DE CONFIANZA

En estadística inferencial, un tema importante es la construcción de intervalos de confianza [9]. Un intervalo de confianza da un rango estimado de valores que es probable que incluya un parámetro poblacional desconocido. Este rango estimado se calcula a partir de un determinado conjunto de datos seleccionados

aleatoriamente de una muestra de una población de interés. La selección del nivel de confianza para el intervalo determina la probabilidad de que el intervalo de confianza construido contendrá al valor verdadero del parámetro. Opciones comunes para el nivel de confianza, "C", son 0.90, 0.95, y 0.99. Estos niveles corresponden a porcentajes del área de la curva de densidad normal. Por ejemplo, un intervalo de confianza de 0.95 cubre el 95% de la curva normal. Esto significa que la probabilidad de observar un valor fuera de esta área es de menos de 0.05 (o 5%). Debido a que la curva normal es simétrica, una mitad del área se encuentra en la cola izquierda de la curva y otra mitad en la cola derecha de la misma. Como se muestra en la figura 1, para un intervalo de confianza con el nivel C, la zona en cada cola de la curva es igual  $\alpha/2 = (1 - C)/2$ . Por ejemplo, para un intervalo de confianza de 0.95, el área en cada cola es igual a  $0.05/2$ , que es igual a 0.025. Para determinar los dos puntos críticos (izquierda y derecha de la cola), se necesitan tablas estadísticas de la distribución normal [10]. En nuestro sistema estas tablas no son necesarias ya que los valores son calculados automáticamente por el software.

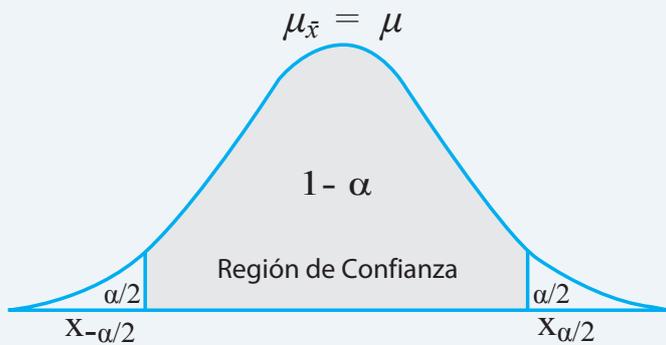


Figura 1. Intervalo de confianza para el promedio de la población.

## RESULTADOS

En esta segunda etapa del proyecto, hemos desarrollado objetos de enseñanza y aprendizaje para la estadística inferencial. Cada objeto calcula un intervalo de confianza para una situación distinta. El primer objeto calcula intervalos de confianza para la media aritmética de la población. La aplicación utiliza la ecuación (1) si se conoce el valor de la desviación estándar de la población y usa la ecuación (2) si no se conoce:

$$\bar{X} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (1)$$

$$\bar{X} \pm z_{1-\alpha/2} \frac{S}{\sqrt{n}} \quad (2)$$

En las ecuaciones (1) y (2)  $\bar{X}$  es la media aritmética de la muestra,  $\sigma$  es la desviación estándar de la población,  $S$  es la desviación estándar de la muestra,  $n$  es el tamaño de la muestra y  $z_{1-\alpha/2}$  es el valor crítico de la distribución estándar normal. La figura 2 muestra la salida para una muestra de 57 datos, con una media de 7.5, desviación estándar de 2.1 y un nivel de confianza de 95%.

El segundo objeto calcula intervalos de confianza para la media de la población cuando el tamaño de la muestra es inferior a 30, la población tiene una distribución normal y la desviación estándar es desconocida. Para este caso, la distribución  $t$  de student es la manera adecuada para construir el intervalo de confianza para la media de la población. En este caso el sistema utiliza la ecuación (3):

$$\bar{X} \pm t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} \quad (3)$$

En la ecuación (3)  $\bar{X}$  es la media de la muestra,  $S$  es la desviación estándar de la muestra,  $n$  es el tamaño de la

muestra,  $n-1$  son los grados de libertad, y  $t_{n-1, 1-\alpha/2}$  es el valor crítico correspondiente para el grado de confianza  $1-\alpha/2$  de la distribución  $t$  student. La figura 3 muestra la salida de una muestra de 21 datos, con promedio de 7.5, desviación estándar de 2.1 y un nivel de confianza de 95%.

El tercer objeto calcula intervalos de confianza para la media de la población cuando la población tiene una distribución no normal. En este caso, se utiliza el teorema de Chebyshev para construir el intervalo de confianza. La aplicación utiliza la ecuación (4) si se conoce la desviación estándar de la población y usa la ecuación (5) si no se conoce:

$$\bar{X} \pm k \frac{\sigma}{\sqrt{n}} \quad (4)$$

$$\bar{X} \pm k \frac{S}{\sqrt{n}} \quad (5)$$

En las ecuaciones (4) y (5)  $\bar{X}$  es el promedio de la muestra,  $\sigma$  es la desviación estándar de la población,  $S$  es la desviación estándar de la muestra,  $n$  es el tamaño de la muestra y  $k = \sqrt{1/\alpha}$ , donde  $\alpha$  es el nivel de significación. La figura 4 muestra la salida para una muestra de datos, con un promedio de 7.5, desviación estándar de 2.1 y un nivel de confianza de 95%.

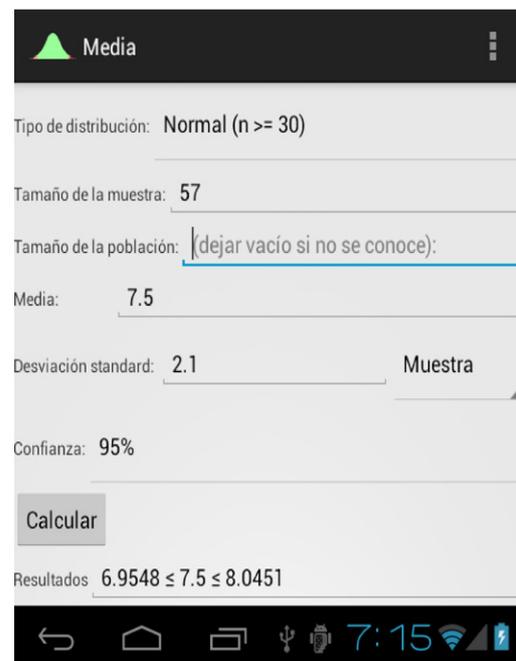


Figura 2. Intervalo de confianza de la media para una distribución normal.

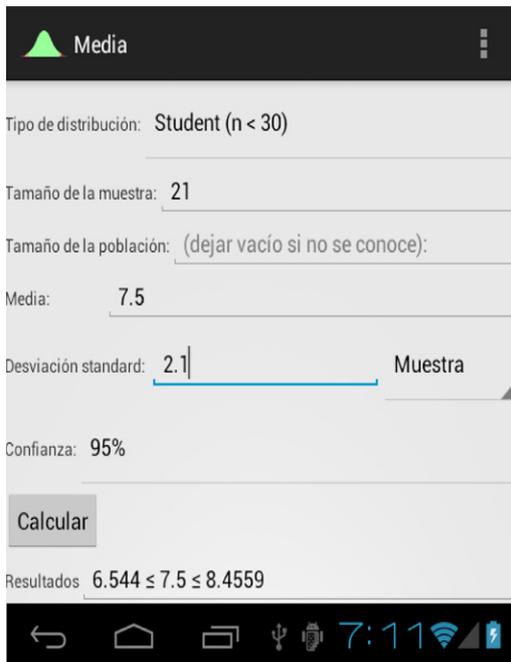


Figura 3. Intervalo de confianza de la media para una distribución de student.

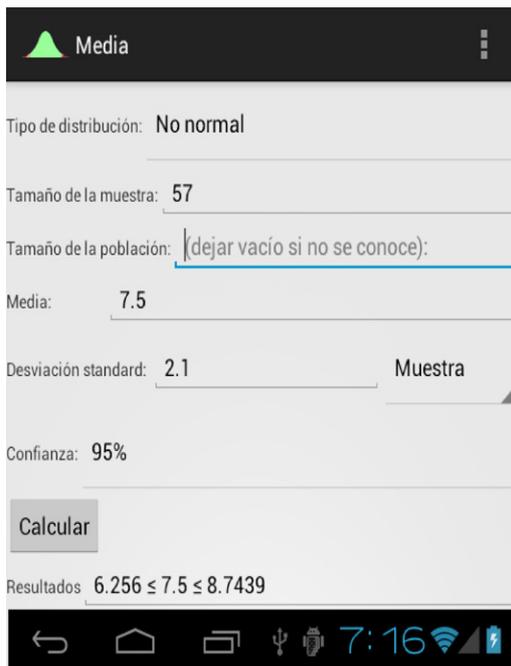


Figura 4. Intervalo de confianza de la media para una distribución no normal.

## CONCLUSIONES

Hemos presentado los principales logros que se han obtenido al finalizar la segunda etapa del proyecto de "Statistics-to-Go", en donde hemos diseñado herramientas de aprendizaje que mejoran el aprendizaje intuitivo de conceptos estadísticos utilizando exclusivamente un

teléfono celular o tableta electrónica y aprovechando la ubicuidad de la tecnología personal con fines educativos, ayudando sobre todo a aquellos estudiantes que no tienen posibilidades para la compra de una computadora o el pago de servicios de Internet. Estas herramientas de aprendizaje cubren la mayor parte del programa de estudios del curso de estadística descriptiva e inferencial introductoria ofrecido en el primer semestre en la Universidad de Sonora.

En la tercera etapa de este proyecto se diseñarán herramientas de enseñanza y aprendizaje para calcular intervalos de confianza para la desviación estándar de la población, el cociente de dos varianzas, la diferencia entre dos medias aritméticas y la diferencia entre dos proporciones. Asimismo, probaremos la eficacia de los objetos diseñados en los dispositivos móviles de los estudiantes (teléfonos celulares y tabletas electrónicas) y recabaremos opiniones de los alumnos sobre el uso de nuestras aplicaciones.

## BIBLIOGRAFÍA

- 1) F. Tapia et al. «Elaboration of Statistics Learning Objects for Mobile Devices,» IJIM, May 2012.
- 2) B. Cope, M. Kalantzis. «Ubiquitous Learning. Exploring the anywhere/anytime possibilities for learning in the age of digital media,» University of Illinois Press. 2009.
- 3) J. Traxler. «Defining mobile learning,» In: international conference mobile learning 2005. 261-266. Malta. 2005.
- 4) Wikipedia, the free encyclopedia. Android (operating system). [En línea]. Available: [http://en.wikipedia.org/wiki/Android\\_\(operating\\_system\)](http://en.wikipedia.org/wiki/Android_(operating_system)). [Último acceso: 20 Junio 2014].
- 5) [En línea]. Available: <http://www.tech-thoughts.net/2012/12/smartphone-market-share-trends-by-country.html#.UQR-SaL-mz2Y>. [Último acceso: 7 Junio 2014].
- 6) [En línea]. Available: [http://www.kantarworldpanel.com/dwl.php?sn=news\\_downloads&id=113](http://www.kantarworldpanel.com/dwl.php?sn=news_downloads&id=113). [Último acceso: 7 Junio 2014].
- 7) [En línea]. Available: <http://www.kantarworldpanel.com/global/News/news-articles/Windows-sees-strong-European-growth>. [Último acceso: 20 Junio 2014].
- 8) [En línea]. Available: <http://bgr.com/2012/12/05/tablet-market-share-2012/>. [Último acceso: 20 Junio 2014].
- 9) M. F. Triola. «Elementary Statistics with Multimedia Study/Guide,» 10th Edition. Addison-Wesley Longman. 2007.
- 10) H. Neave. «Statistics Tables: For Mathematicians, Engineers, Economists and the Behavioral and Management Sciences,» College Audience. 2011.

