

# ¡Estadístic@s en acción!

ESTADÍSTICA

PROBABILIDAD Y ESTADÍSTICA II  
**2015**

Soledad DUPRAT - Joana PIZARRO MOLINA

Williams URIBE - Priscila IRRIBARRA

Daira BERNÁRDEZ - Julieta CURTO

Sebastián CERANI OROSCO - Manuel OLIVERA

Daniela DASSO - Déborah MENDOZA VIRGILI

Micaela BELUZZO - Nerina BRAUN

Compiladoras

Maria Paula DIESER - Lorena Veronica CAVERO



FACULTAD DE CIENCIAS  
EXACTAS Y NATURALES

Universidad Nacional de La Pampa

# Índice

<b>Prólogo</b>	<b>2</b>
<b>Variación del pulso después de hacer ejercicio</b> <i>Soledad Duprat y Joana Pizarro Molina</i>	<b>4</b>
<b>¿Cuánto mide el radio de la Tierra?</b> <i>Williams Uribe y Priscila Irribarra</i>	<b>11</b>
<b>Efecto de lesiones cerebrales</b> <i>Daira Bernárdez y Julieta Curto</i>	<b>18</b>
<b>Ley de Benford</b> <i>Sebastián Cerani Orosco y Manuel Olivera</i>	<b>25</b>
<b>Calidad de los generadores de números pseudoaleatorios</b> <i>Daniela Dasso y Déborah Mendoza Virgili</i>	<b>31</b>
<b>Objetos del diseño ergonómico</b> <i>Micaela Beluzzo y Nerina Braun</i>	<b>40</b>

## Prólogo

Hace ya algunos años, como respuesta a las demandas de nuevas competencias profesionales por parte de una sociedad basada en la economía de servicios, y debido a las investigaciones sobre la manera de enseñar y aprender de las personas, se piensa en una educación eficaz en la medida que ésta sea capaz de desarrollar habilidades de alto nivel que ayuden a los estudiantes a aprender a lo largo de toda su vida, *i.e.* una educación capaz de ofrecer a los ciudadanos un conocimiento sólido y a la vez flexible que pueda dar respuestas ajustadas a las situaciones cambiantes que se presentan (Barberà, 2005). La enseñanza de la estadística no ha quedado ajena a esta perspectiva de transformación, siendo objeto de un marcado interés en la comunidad educativa, en donde se habla de construir una cultura estadística. Gal (2002) se refiere a dos componentes interrelacionados: la capacidad para interpretar y evaluar críticamente la información estadística, y la capacidad para discutir o comunicar sus opiniones respecto a tal información. Franklin *et al.* (2005) indican que la enseñanza de la estadística debe, como principal objetivo, ayudar a los estudiantes a aprender los elementos básicos del pensamiento estadístico como la importancia de los datos, la ubicuidad de la variabilidad, así como su cuantificación y explicación. Según Wild & Pfannkuch (1999), el razonamiento estadístico es esencial para el aprendizaje e incluye cinco componentes fundamentales: reconocer la necesidad de los datos, la transnumeración, percibir la variabilidad, razonar con modelos estadísticos e integrar la estadística al contexto.

Asimismo, es necesario diferenciar entre conocer y ser capaz de aplicar un conocimiento. La habilidad para aplicar los conocimientos estadísticos requiere no sólo de conocimientos técnicos (construir un gráfico o calcular un promedio), sino también conocimientos estratégicos (saber cuándo usar un concepto o gráfico dado). Los problemas y ejercicios presentes en libros de texto sólo se concentran en conocimientos técnicos. Sin embargo, el trabajo con datos reales demanda utilizar conocimientos estratégicos.

En consecuencia, suponemos que la mejor forma de seguir estas recomendaciones es introducir, en las clases de estadística, problemas con datos reales, sean propuestos por el profesor o diseñados por los estudiantes. En este tipo de experiencias, se procura reemplazar la introducción de conceptos y técnicas descontextualizadas, o aplicadas a problemas tipo, difíciles de encontrar en la vida real, por una actividad integral donde se presenten y desarrollen las diferentes fases de una investigación estadística. Por estos motivos, desde 2011, proponemos a nuestros estudiantes de las asignaturas “Estadística” y “Probabilidad y Estadística II” de las carreras de Profesorado y Licenciatura en Matemática de la Facultad de Ciencias Exactas y Naturales (UNLPam), respectivamente, la realización de proyectos con datos reales desarrollando las diferentes etapas de una investigación

estadística, con los objetivos no sólo de promover el aprendizaje de la estadística a partir del trabajo con datos reales, fomentando el uso de las TIC como herramientas necesarias en la resolución de este tipo de problemas, sino además estimular una mayor reflexión, compromiso y proactividad.

En este documento reunimos los trabajos de nuestros estudiantes del Ciclo Lectivo 2015, quienes abordaron temáticas de lo más diversas: leyes probabilísticas vinculadas con fenómenos corrientes, normas en diseños de mobiliario, investigaciones médicas, entre otras.

*Las profes, noviembre de 2015*

---

## Referencias

- Barberá, E. (2005). La evaluación de competencias complejas: la práctica del portafolio. *Educere*, 31(9), pp 497-504.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association. Disponible en: [www.amstat.org/Education/gaise/](http://www.amstat.org/Education/gaise/)
- Gal, I (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Wild, C. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-265.

## VARIACIÓN DEL PULSO DESPUÉS DE HACER EJERCICIO

Duprat, Carmen Soledad – Pizarro Molina, Joana Yanela  
[soledadduprat@gmail.com](mailto:soledadduprat@gmail.com) – [joanapizarro.9@gmail.com](mailto:joanapizarro.9@gmail.com)

### Resumen

En el presente trabajo se encuentra una investigación sobre la variación de los pulsos de personas que realizaron diferentes actividades (correr o sentarse) para determinar si existen diferencias entre estos. Para ello, se realizó la comparación de las variaciones de los pulsos arteriales teniendo en cuenta la actividad realizada con el objetivo de establecer si dicha variación de pulsos de las personas que optan por correr es mayor respecto de las que deciden sentarse, dichas mediciones se realizaron en 109 personas de entre 18 y 45 años de edad.

El análisis de la variación de los pulsos nos permitirá observar el funcionamiento del cuerpo, principalmente del corazón y cómo actúa éste ante las distintas acciones.

### Introducción

El análisis de la variación de los pulsos, entre personas que deciden correr o sentarse, es importante porque nos permite saber si, al compararlas, existen diferencias entre los pulsos arteriales de las mismas al realizar las diferentes actividades; por otro lado, permite entender el funcionamiento del corazón y observar cómo se comporta ante las distintas acciones.

Teniendo en cuenta que el pulso forma parte de los signos vitales, los cuales son las señales o reacciones que presenta un ser humano con vida, que revelan las funciones básicas del organismo, varían de individuo a individuo según la edad y en diferentes horas del día y a la diferente actividad; resulta significativo observar la variación del mismo para poder detectar el funcionamiento del cuerpo (González, 2009).

Para esto sería conveniente conocer que el Pulso Arterial es la onda pulsátil de la sangre, originada en la contracción del ventrículo izquierdo del corazón y que resulta en la expansión y contracción regular del calibre de las arterias.

La onda pulsátil representa el rendimiento del latido cardíaco, que es la cantidad de sangre que entra en las arterias con cada contracción ventricular y la adaptación de las arterias, o sea, su capacidad de contraerse y dilatarse. Asimismo, proporciona información sobre el funcionamiento de la válvula aórtica.

El pulso periférico se palpa fácilmente en pies, manos, cara y cuello. Realmente puede palparse en cualquier zona donde una arteria superficial pueda ser fácilmente comprimida contra una superficie ósea.

La velocidad del pulso (latidos por minuto) corresponde a la frecuencia cardíaca, la cual varía con la edad, sexo, actividad física, estado emocional, fiebre, medicamentos y hemorragias (Penagos, S. P., Salazar, L. D., Vera, F. E., 2005, p 1466-1467).

En la Tabla 1 podemos observar los valores normales de frecuencia cardíaca, donde se indica el rango medio considerado normal de pulsaciones por minuto teniendo en cuenta la edad de la persona:

EDAD	PULSACIONES POR MINUTO
Recién nacido	120 – 170
Lactante menor	120 – 160
Lactante mayor	110 – 130
Niños de 2 a 4 años	100 – 120
Niños de 6 a 8 años	100 – 115
Adulto	60 – 80

Tabla 1. Valores Normales de Frecuencia Cardíaca

El objetivo de este trabajo es analizar la variación del pulso, en personas de entre 18 y 45 años de edad a las cuales se le asigna una determinada actividad (correr, sentarse), para observar si es mayor en las personas que eligen correr que en las personas que se sientan.

## Metodología

El estudio se realizó sobre una muestra aleatoria simple, obtenida por el Prof. John Eccleston y el Dr. Richard Wilson de la Universidad de Queens-land, de 110 (de los cuales trabajamos con 109 porque la persona número 76 no aportaba información) estudiantes de entre 18 y 65 años de edad, a los que se les pidió que se midieran el pulso arterial antes y después de realizar la actividad asignada.

Los estudiantes de cinco grupos de clases entre 1993 y 1998 participaron en un experimento simple que consistió en tomar su propio pulso. Se les pidió que lanzaran una moneda:

- Si salía cara: tenían que correr en el lugar durante un minuto.
- Caso contrario: se sentaban durante un minuto.

Tras el minuto cada uno tomó su pulso de nuevo.

Como al Prof. Wilson le preocupaba que algunos estudiantes registraran la opción menos exigente en lugar de correr, incluso si la moneda mostraba cara, decidió modificar el método para el periodo de 1995 a 1998; el cual consistió en entregar formularios para consignar los datos antes de realizar el experimento, donde las opciones fueron preasignadas en la misma cantidad (sentarse o correr).

Se efectuó un análisis descriptivo sobre la variación de los pulsos (obtenida al realizar la diferencia de los pulsos, después y antes) teniendo en cuenta la actividad y se realizaron inferencias sobre ella utilizando el software estadístico R a través de su interfaz gráfica R Commander (R Core Team, 2015).

Para llevar a cabo el objetivo de estudio primero se utilizó una prueba chi-cuadrado de bondad de ajuste con un nivel de 0.05 para poner a prueba la normalidad de las variables estudiadas.

Además se realizaron test de hipótesis e intervalos de confianza de nivel 0.05 para las medias y varianzas de cada una de las poblaciones estudiadas; teniendo en cuenta que fue probada la normalidad podemos especificar que lo que hemos utilizado es Test de hipótesis para el cociente de varianzas de poblaciones normales.

Por último para observar si la variación del pulso es mayor en las personas que eligen correr se realizó un test de hipótesis para la diferencia de medias de poblaciones normales (muestras aleatorias independientes con varianzas desconocidas y distintas) con un nivel de 0.05.

## Resultados

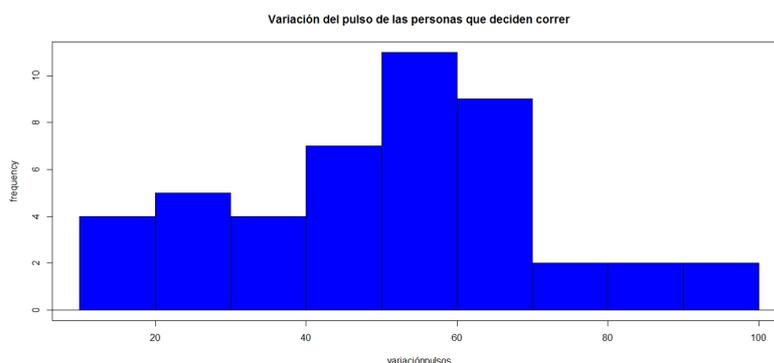
A continuación se muestran los resultados obtenidos en el análisis de cada una de las variables, como así también los resultados obtenidos en la comparación realizada sobre las mismas.

### Análisis de los datos de las personas que deciden correr:

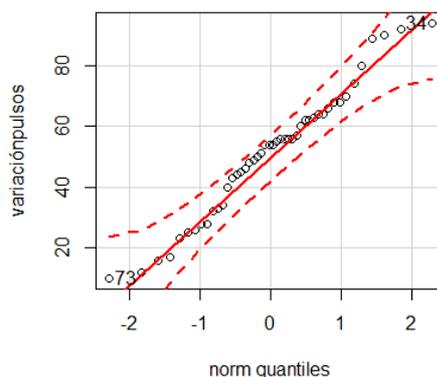
La variación del pulso de las personas entre 18 y 45 años de edad que decidieron correr es  $51.3913 \pm 21.08973$  (CV= 0.41038, lo que indica que la dispersión de los datos es alta).

En la Figura 1 podemos observar que:

Con mayor frecuencia, los estudiantes que deciden correr poseen una variación de pulso entre 50 y 60 ppm.



**Figura 1:** Histograma de la variación del pulso de las personas que deciden correr.



**Figura 2:** QQ-plot de la variación del pulso de las personas que deciden correr.

Analizando la Figura 2, se podría suponer que los datos provienen de una distribución normal, ya que los puntos (determinados por los valores observados y esperados bajo normalidad) se acercan a una recta.

Análogamente, observado las Figuras 1 y 2, podemos suponer que la variable en estudio sigue una distribución normal cuyos parámetros estimados, por el método de máxima verosimilitud, son:

$$\mu_1^* = 51.3913$$

$$\sigma_1^* = 21.08973$$

Cabe aclarar, que dado que los parámetros poblacionales son desconocidos, se estimaron a partir de la muestra, siendo el estimador del primer parámetro la media muestral ( $\mu_1^* = \bar{x}$ ) y la desviación estándar ( $\sigma_1^* = s_1$ ) para el segundo. Los estimadores utilizados son insesgados, eficientes y suficientes (IMVU).

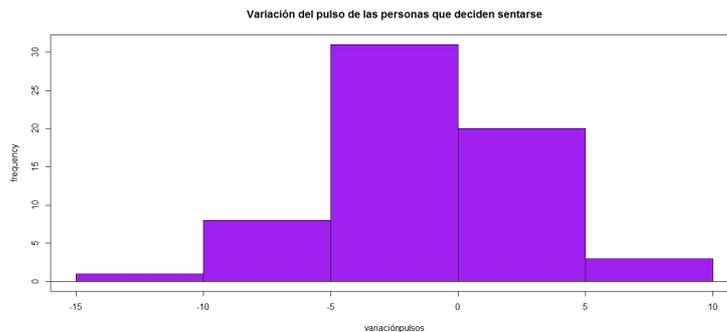
Con una prueba chi-cuadrado de bondad de ajuste pudimos corroborar que efectivamente la variable en estudio sigue una distribución normal. (chi-cuadrado= 3.6971, g.l= 3, p=0.295734).

#### Análisis de los datos de las personas que deciden sentarse:

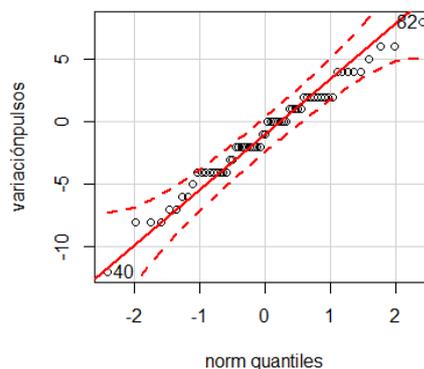
La variación del pulso de las personas entre 18 años y 45 años de edad que decidieron sentarse es  $-1 \pm 3.922639$  ppm ( $|CV| = 3.922639$ , en este caso el CV indica que la dispersión es muy alta).

En la Figura 3 se puede observar que:

Con mayor frecuencia, los estudiantes que deciden sentarse poseen una variación de pulso entre -5 y 0 ppm.



**Figura 3:** Histograma de la variación del pulso de las personas que deciden sentarse.



**Figura 4:** QQ-plot de la variación del pulso de las personas que deciden sentarse.

Respecto del gráfico de la Figura 4, también se observa que los puntos se acercan a una recta y sobre todo en la zona central, por lo que podría suponerse que la distribución de los datos es normal. Finalmente, las Figuras 3 y 4 nos sugieren que la variable en estudio sigue una distribución normal cuyos parámetros estimados, por el método máxima verosimilitud, son:

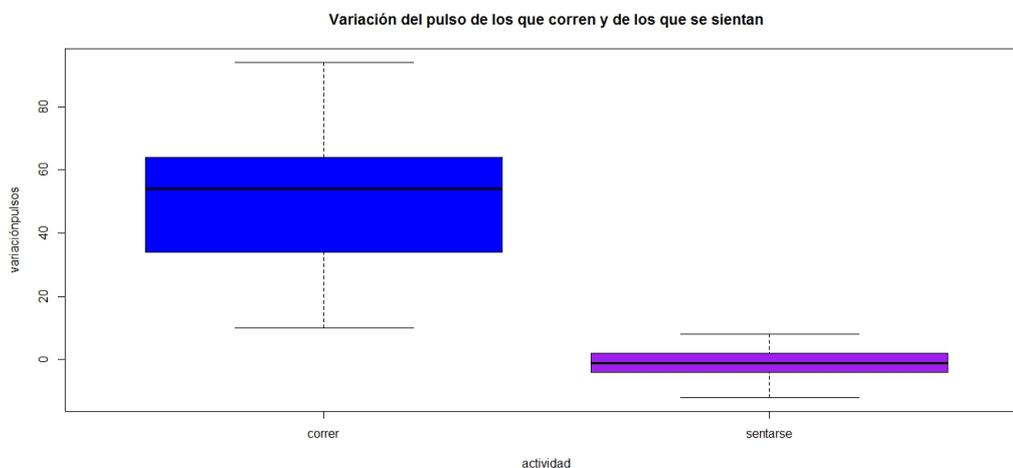
$$\mu_2^* = -1$$

$$\sigma_2^* = 3.922639$$

En este caso también ocurre que, como los parámetros poblacionales son desconocidos, se estimaron a partir de la muestra, siendo el estimador del primer parámetro la media muestral ( $\mu_2^* = \bar{y}$ ) y la desviación estándar ( $\sigma_2^* = s_2$ ) para el segundo. Los estimadores utilizados son insesgados, eficientes y suficientes (IMVU).

Con una prueba chi-cuadrado de bondad de ajuste pudimos corroborar que la variable en estudio sigue una distribución normal (chi-cuadrado=0.40695, g.l= 3, p=0.9951).

Análisis de los datos de las personas que deciden correr conjuntamente con las que deciden sentarse:



**Figura 5:** Boxplot de la variación del pulso de las personas que corren y de las personas que se sientan.

Respecto de la Figura 5 podemos observar que la variación del pulso en personas que deciden correr presenta una mayor dispersión con respecto al de los que deciden sentarse, también es posible ver que las medianas de dichos gráficos se ubican aproximadamente en los centros de las cajas y que entre ellas son diferentes; en ambos casos se presenta una distribución simétrica de la variación de los pulsos.

Con el propósito de analizar si las varianzas de las variables analizadas son distintas realizamos la prueba F para el cociente de varianzas (donde  $H_0 : \frac{\sigma_1}{\sigma_2} \leq 1$  vs  $H_a : \frac{\sigma_1}{\sigma_2} > 1$  y se obtuvo:  $F = 28.906$ , num df = 45, denom df = 62, p-value < 2.2e-16 cuyo IC= (16.89852; 50.65493)), con lo cual pudimos afirmar que efectivamente era de esta manera y además que  $\sigma_1 > \sigma_2$ .

Para determinar si la variación de pulsos en personas de entre 18 y 45 años de edad que deciden correr es mayor al de los que deciden sentarse; se realizó un test de hipótesis sobre la diferencia de medias para poblaciones normales teniendo en cuenta la actividad realizada, cuyas muestras aleatorias son independientes con varianzas desconocidas y distintas al nivel 0.05. Luego de haberlo

realizado pudimos concluir, mediante el test de hipótesis unilateral ( $H_0 : \mu_1 - \mu_2 \leq 0$  vs  $H_a : \mu_1 - \mu_2 > 0$ ;  $t= 16.64$ ,  $df= 47.28$ ,  $p\text{-value} < 2.2e-16$ ), que efectivamente hay suficiente evidencia para suponer que la media de la variación de los pulsos de las personas que deciden correr es mayor que la media de la variación de los pulsos de las personas que deciden sentarse.

Ahora bien, teniendo en cuenta el IC=(46.05825; 58.72436) obtenido en este caso, podemos decir que en promedio la diferencia de medias es de  $52.391305 \pm 6.33305$  aproximadamente.

## Discusión

Analizando los datos obtenidos de la variación de los pulsos de las personas teniendo en cuenta la actividad realizada, tenemos que:

	Personas que deciden correr	Personas que deciden sentarse
Número de personas	46	63
La variación de los pulsos se encuentran concentrados entre	50 y 60 ppm	-5 y 0 ppm
Teniendo en cuenta los Histogramas (Figura 1 y 3) realizados concluimos que	Aparentan simetría respecto de los valores antes mencionados (se observan las clases más frecuentes localizadas a un lado y a otro de dicha medida, mientras que las menos frecuentes se encuentran en los extremos)	
Analizando los boxplots (Figura 5) podemos afirmar que se observa	Mayor dispersión en la variación de los pulsos de las personas que deciden correr respecto de las que optan por sentarse. Ambos gráficos muestran una distribución simétrica de las variaciones de los pulsos teniendo en cuenta que las medianas en los mismos se ubican aproximadamente en los centros de cada uno.	
Supuestos	Las variables siguen una distribución normal.	

Tabla 2. Resumen del análisis de los gráficos.

Por otro lado las pruebas chi-cuadrado de bondad de ajuste arrojaron que las variables se distribuyen normalmente con los parámetros establecidos. Esto nos permitió realizar el test de hipótesis sobre la diferencia de medias correspondiente, mediante el cual determinamos que la media de la variación de los pulsos de las personas que deciden correr es mayor que la media de la variación de los pulsos de las personas que deciden sentarse.

Entonces, por lo antes mencionado y observando la Tabla 2, podemos afirmar que la variación de los pulsos de las personas que deciden correr es mayor a la de las que deciden sentarse. Parecería ser que el corazón de las personas que hacen actividad física (caminar, correr, bailar, etc) late con mayor velocidad que el de las personas que permanecen quietas.

## Conclusiones

Teniendo en cuenta que el objetivo de este trabajo se centró en analizar la variación de pulsos de las personas que corren y se sientan; que nos interesa observar si la variación de las primeras era mayor a las de las segundas (nuestra hipótesis) y los resultados obtenidos luego de los análisis realizados, descriptivo e inferencial, podemos decir que logramos corroborar que la variación del pulso de las personas que decidieron correr es mayor que en las personas que decidieron sentarse.

Por otro lado, nos parecería interesante realizar a futuro un análisis más profundo de este tema; ya que solo en este caso se estudió la variación del pulso teniendo en cuenta la actividad realizada pero podríamos analizar lo mismo teniendo en cuenta factores como: la frecuencia con la que la persona realiza ejercicio, el sexo, si posee hábitos como fumar o ingerir alcohol, etc. Este análisis detallado nos permitiría observar la diferencia de variación de pulsos de manera más exacta llegando a mejores conclusiones sobre esta problemática.

## Referencias

- González Maita, R. A. (2009). *Técnica de la toma de signos vitales en el adulto* (Doctoral dissertation). Disponible en: <http://dspace.unl.edu.ec/jspui/bitstream/123456789/4492/1/GONZ%C3%81LEZ%20MAITA%20ROSA%20ALEJANDRINA.pdf> (12 de Noviembre de 2015).
- Penagos, S. P., de Urgencias, E., Salazar, L. D., de Enfermería, E. C., Vera, F. E., & Cardioinfantil, F. (2005). Control de signos vitales. *Guías para manejo de Urgencias. Bogotá, Colombia: Fundación Cardioinfantil*, p 1465 a 1473. Disponible en: [http://biblioises.com.ar/Contenido/600/612/Control\\_de\\_signos\\_vitales.pdf](http://biblioises.com.ar/Contenido/600/612/Control_de_signos_vitales.pdf) (12 de Noviembre de 2015).
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Versión 3.2.2 (2015-08-14)) [Software]. Disponible en: <http://www.Rproject.org/> (12 de Noviembre de 2015).



## ¿CUÁNTO MIDE EL RADIO DE LA TIERRA?

Williams Noel Uribe - Ana Priscila Irribarra

[uribewilliams@gmail.com](mailto:uribewilliams@gmail.com) - [priscila.irribarra.pi@gmail.com](mailto:priscila.irribarra.pi@gmail.com)

### Resumen

Hace un poco más de 2.246 años, Eratóstenes, matemático, astrónomo y geógrafo griego, trató de dar respuesta a esta pregunta. Para ello, en la ciudad egipcia de Alejandría, Eratóstenes midió el ángulo de la sombra que proyectaba un palo clavado en el suelo durante el mediodía solar. Exactamente un año después, a la misma hora, comprobó en la ciudad de Assuan que el palo no daba ninguna sombra. Luego calculó la distancia que separaba a ambas ciudades egipcias. Hizo un cálculo matemático, y así determinó el diámetro de la Tierra y su radio.

En este trabajo se presenta un análisis de 85 valores del radio de la Tierra calculados siguiendo el método de Eratóstenes. Para ello se realizó un estudio descriptivo e inferencial para determinar si estos valores difieren del aceptado hoy en día.

### Introducción

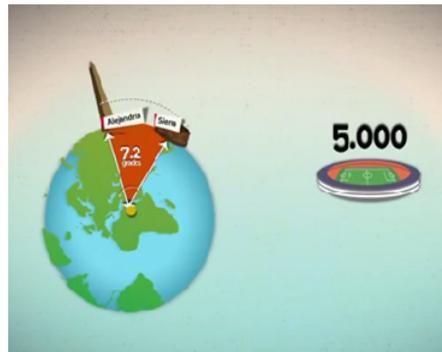
Eratóstenes nació en Cirene en el año 276 a.C. y se cree que era de origen caldeo. Fue matemático, astrónomo y geógrafo. Alrededor del año 255 a.C. fue nombrado director de la Biblioteca de Alejandría por el rey Ptolomeo. Trabajó con problemas de matemáticas, como la duplicación del cubo y los números primos. Además de matemático, Eratóstenes era astrónomo y una de sus principales contribuciones a la ciencia y a la astronomía fue su trabajo sobre la medición de la Tierra. Estando en la Biblioteca de Alejandría, encontró un informe de observaciones sobre Siena, ciudad situada a unos 800 km al sur de Alejandría, en el que se decía que el día del solsticio de verano (21 de junio) a mediodía, los objetos no producían sombra y en el fondo de los pozos podía verse la luz del sol. Pero observó que, en Alejandría, el mismo día y a la misma hora no se producía este mismo hecho. Asumió de manera correcta que el Sol se encontraba a gran distancia y que sus rayos, al alcanzar la tierra, lo hacían en forma (prácticamente) paralela (Gómez Aroca, 2013).

Esto ratificaba su idea de que la superficie de la Tierra era curva pues, de haber sido plana, no se hubiese producido esta diferencia entre las dos ciudades. El siguiente paso fue medir en Alejandría el ángulo que formaban los rayos del sol con la vertical que, por construcción, es igual al ángulo cuyo vértice está en el centro de la Tierra.

Como podemos observar en la Figura 1, dicho ángulo fue de  $7,2^\circ$  ¿pero cómo obtuvo ese valor? Hoy en día, entre los historiadores de las ciencias existen dos teorías que gozan de gran aceptación: una es que Eratóstenes calculó la razón entre la longitud de la sombra proyectada y la de la columna vertical, encontrando que era de  $\frac{1}{8}$ , con lo cual utilizando unas tablas de equivalencias con tangente le permitió encontrar el ángulo buscado. La otra teoría es que utilizó dos instrumentos de la época para su medición: el Gnomon<sup>1</sup> y el Scaphe<sup>2</sup>, que utilizados juntos le permitieron calcular

<sup>1</sup>Un objeto alargado cuya sombra se proyectaba sobre una escala graduada para medir el paso del tiempo.

el ángulo que forma la sombra con la hipotenusa, luego utilizando el Teorema de la suma de los ángulos interiores de un triángulo, obtuvo el ángulo buscado. Cualquiera hubiera sido el método lo importante era hallar el ángulo.



**Figura 1:**Ángulo formado entre Alejandría y Siena, con vértice en el centro de la Tierra.

Conociendo este ángulo y la distancia entre las ciudades, que era de 5000 estadios<sup>3</sup> obtuvo que la circunferencia de la tierra media  $\frac{360 \times 5000}{7.2} = 250000$  estadios, aproximadamente 40000 km. Lo que equivale a un radio de 6366 km, con un error bastante pequeño, contra los 6371 km (Sáez Pastor, 2010) que son los admitidos hoy en día, lo que constituye un logro notable para el año en que tuvo lugar.

El Proyecto Eratóstenes es una propuesta del Departamento de Física de la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires (Argentina), del Laboratorio Pierre Auger, Universidad Tecnológica Nacional, Regional Mendoza (Argentina) y de la Asociación Física Argentina (Carbajal, 2012).

Los objetivos del proyecto, básicamente son:

- Describir cómo el perímetro y radio de la Tierra fue medido por primera vez miles de años atrás.
- Calcular el radio terrestre con el aporte de varios grupos que forman parte del proyecto y comparar si efectivamente el radio calculado difiere del calculado científicamente.

En sí, es una adaptación para la Argentina del proyecto *WYP Eratosthenes Project* organizado en Estados Unidos en ocasión del Año Internacional de la Física 2005.

También en 2009, al cumplirse el cuarto centenario del uso del primer telescopio para la observación astronómica por parte de Galileo, una asociación de Barcelona propuso a los centros educativos el proyecto de medir el radio de la Tierra según el método de Eratóstenes en la antigüedad.

### Metodología

Para este trabajo se consideró como población de estudio a todas las mediciones del radio de la Tierra en las que se reproduce el método utilizado por Eratóstenes para calcularlo. Sobre esta población se consideró una muestra de 85 valores del radio de la Tierra, obtenidos por todos los pares de instituciones inscriptas en el Proyecto Eratóstenes en el año 2012. Proyecto en el cual

<sup>2</sup>Instrumento griego que permite saber la hora solar y cuándo es el solsticio de verano o el de invierno.

<sup>3</sup>El estadio era una unidad de longitud griega, que tomaba como patrón la longitud del estadio de Olimpia, que equivalía a 174,125 metros. El estadio que empleó Eratóstenes para medir la circunferencia polar de la Tierra, se estima que era aproximadamente de 185 metros (estadio egipcio)

participaron más de 12000 alumnos de 167 escuelas americanas y dos españolas. Las mediciones se pautaron entre el 10 y el 21 de septiembre. Como cada cálculo del radio terrestre demanda, al menos, dos escuelas que midan sombras y longitudes de “gnomones”, se formaron 110 pares de escuelas distribuidas de acuerdo a métodos de optimización en base a las coordenadas geográficas de todos los participantes y al día de medición. Con estos métodos, los referentes del proyecto tienen en cuenta algunos aspectos, que mencionan en la Guía para el Estudiante disponible en su página web, como por ejemplo que las dos escuelas se encuentren sobre un mismo meridiano, separadas a una distancia  $d$  en la dirección norte-sur. Si esto no es así, la distancia  $d$  no debe ser la distancia entre escuelas sino sólo la distancia norte-sur entre ellas. Una manera sencilla de determinar esta última es conocer la distancia entre cada escuela y el Ecuador, distancia que es provista por los organizadores en el momento de la inscripción de la escuela.

Ahora bien, siendo  $X$  = “valor calculado para el radio de la Tierra” nuestra variable de interés, se llevó a cabo un análisis descriptivo e inferencial de la muestra, donde utilizamos R a través de su interfaz gráfica R-Commander (R i386 3.1.1).

Algunas preguntas que guiaron nuestro trabajo fueron: ¿Qué distribución tienen nuestros datos? ¿La medición promedio  $\bar{x}$  difiere del valor aceptado hoy en día?.

Para dar respuesta a nuestro primer interrogante, realizamos un boxplot y al tener una visión de la distribución de los datos, se supuso que seguían una distribución normal, obteniendo la estimación de los parámetros  $\mu$  y  $\sigma$  por máxima verosimilitud. Supuesto que probamos utilizando una prueba chi-cuadrado de bondad de ajuste y presentamos en una misma gráfica el histograma asociado a los datos y la densidad del modelo de ajuste. Esto permite valorar, al menos de forma visual, la precisión del ajuste.

En cuanto al segundo interrogante, para darle respuesta, se puso a prueba que el valor del radio de la tierra corresponde al aceptado a partir de los datos obtenidos de la experiencia mediante un Test de Hipótesis para la media con varianza desconocida de una población normal.

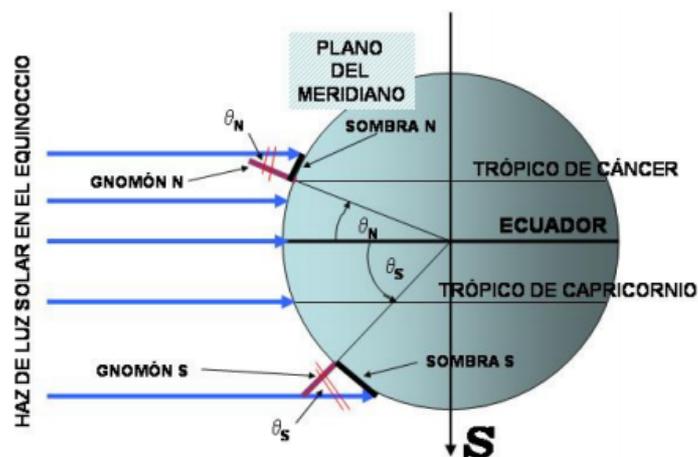
## Resultados

En un principio, fueron 110 las mediciones realizadas en el año 2012 por escuelas de distintos lugares sumados al “Proyecto Eratóstenes” para determinar el radio de la tierra desde sus ciudades. Al analizar esta primera muestra en R y realizar un boxplot, el software marcó 25 datos como atípicos.

Sobre estos 25 datos, tenemos algunas consideraciones:

- Algunos de ellos seguramente están mal tipeados, o corresponden a un resultado previo al cálculo del radio de la Tierra, en este caso el valor obtenido se “aproxima” a su perímetro.
- El resultado obtenido tiene que ver con mediciones tomadas por pares de escuelas cuya distancia entre ellas es pequeña, haciendo que el ángulo entre las dos escuelas, con vértice en el centro de la Tierra sea pequeño y aumenta el error de la medición.
- Proviene de escuelas que no se encuentran sobre el mismo meridiano. En este caso nos preguntamos cuán bien alineadas en la dirección Norte-Sur están, observando que para algunas de las mediciones que se detectaron como atípicas la alineación no es tan buena y en este caso no sabemos si se utilizó (para el cálculo) la distancia entre las dos escuelas o la distancia Norte-Sur entre ellas.
- Proviene de escuelas en hemisferios diferentes. Cuando las escuelas pertenecen a hemisferios diferentes, se tiene que tener el recaudo de que los ángulos tienen signos

opuestos en ambos hemisferios, y el ángulo que se utiliza en el cálculo es el que resulta de sumar los módulos de dos ángulos, cada ángulo formado por una institución y el Ecuador, con vértice en el centro de la Tierra, cómo podemos observar en la Figura 2.



**Figura 2:** Ángulo entre dos escuelas en hemisferios diferentes.

Solo a modo de ejemplo, uno de los valores atípicos registrado es 64140,9 km para el radio de la Tierra. Esta medición la realizaron dos escuelas, una A con coordenadas  $34^{\circ} 35' S - 58^{\circ} 26' O$  y la otra B con ubicación  $26^{\circ} 52' S - 60^{\circ} 13' O$ , ubicaciones que podemos apreciar en el siguiente mapa:

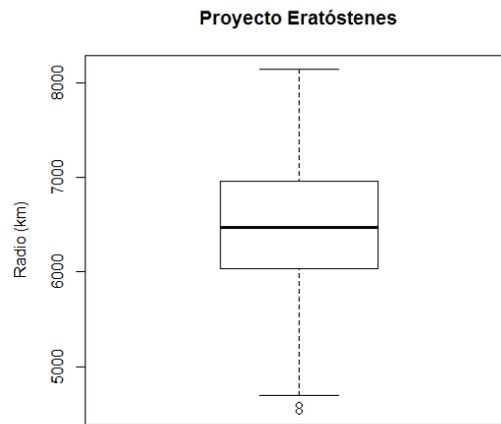


**Figura 3:** Ubicación de dos instituciones cuya medición se consideró atípica

En este ejemplo, las dos instituciones pertenecen al mismo hemisferio (el Sur) y podemos considerar que están alineadas sobre un mismo meridiano. En este caso, creemos que puede haber ocurrido un error de tipeo o que pertenece a un cálculo parcial, en el cual solo se obtuvo el perímetro de la Tierra. Pero no tenemos la información suficiente para descartar que el valor atípico se debe al valor del ángulo obtenido o la distancia entre escuelas utilizada para el cálculo, aunque nos inclinamos más por el primer supuesto.

Por lo antes expuesto, nuestra muestra quedó conformada por 85 mediciones, y para la presentación de los datos y por el tipo de variable a analizar -cuantitativa continua- consideramos oportuno

realizar un boxplot, además de que permite detectar automáticamente cuáles son los datos atípicos y tener un panorama de la distribución de los datos.



**Figura 4:** Boxplot con las mediciones del radio de la Tierra (2012).

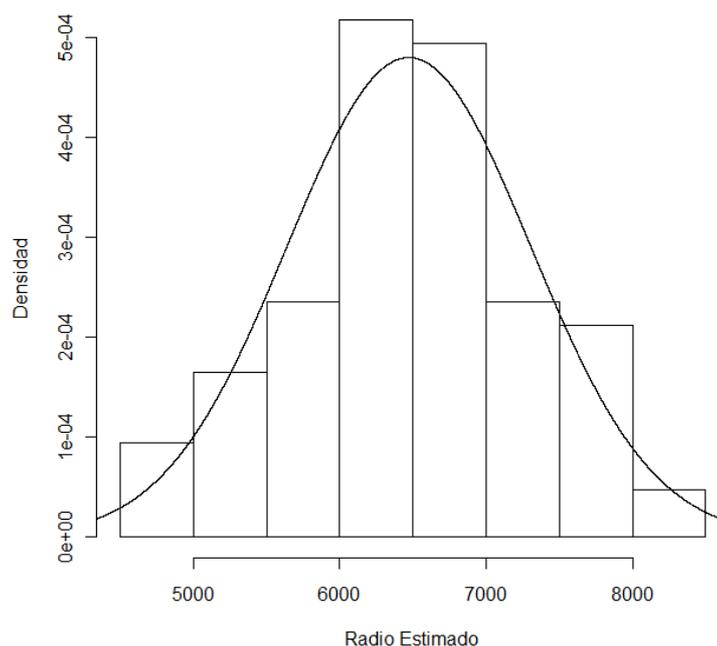
Además luego de un análisis numérico de nuestra variable de interés se obtuvieron los siguientes resultados:

Medida	Media	Mediana	MAD	Desvío	Coefficiente de Variación
<b>Valor obtenido</b>	6472,069 km	6476,06 km	719,0313	831,562 km	0,129

**Tabla 1:** Medidas de tendencia central y de dispersión obtenidas.

Al momento de tener una visión del ajuste, se realizó el histograma de la Figura 4 y sobre él la curva de ajuste de la función de densidad Normal con los valores obtenidos de la muestra para la media y el desvío, valores utilizados como estimadores de los parámetros de una distribución Normal.

En el boxplot de la Figura 4 se proporciona una visión general de la simetría de la distribución de los datos. Además con los resultados expuestos en la Tabla 1, se puede concluir que en promedio el radio de la tierra mide  $\bar{x} = 6472,069$  km, valor que no difiere del valor central de nuestro conjunto de datos, 6476,03 km, característica propia de una distribución simétrica. También podemos decir que las mediciones se desvían en promedio unos 831,562 km respecto de la media. En cuanto a la dispersión de los datos y con los resultados expresados podemos concluir que los datos tienen poca dispersión debido a que el coeficiente de variación  $\frac{s}{\bar{x}} = 0,129$ .



**Figura 5:** Comparación de Histograma, correspondiente a las 85 mediciones, con la curva de densidad ajustada.

Al analizar el boxplot y las medidas de resumen, se puede suponer que los datos siguen una distribución Normal, que se distribuyen en forma simétrica respecto de un eje -valor central de nuestro conjunto de datos- las clases más frecuentes se localizan a uno y a otro lado de la mediana, mientras que los menos frecuentes se encuentran en los extremos. Además al realizar los cálculos correspondientes el 62% de los datos pertenecen al intervalo  $(\bar{x} \pm s)$ , un 95% al  $(\bar{x} \pm 2s)$  y a  $(\bar{x} \pm 3s)$  el 98% de los datos, siendo  $\bar{x} = \hat{\mu}$  y  $s = \hat{\sigma}$  los parámetros estimados de la variable X que suponemos con distribución Normal. El método utilizado para obtener estos estimadores fue el de máxima verosimilitud, garantizando que sean insesgados, suficientes, eficientes y de mínima varianza.

## Discusión

Con los resultados de la regla empírica para diagnósticos de normalidad podemos suponer que la variable en cuestión sigue una distribución Normal. Supuesto, hipótesis nula, que pusimos a prueba mediante un Test chi-cuadrado de bondad de ajuste, con un nivel de significación de  $\alpha = 0,05$ . Al aplicar el test el estadístico de prueba resultó ser *chi – cuadrado* = 2,76, que ha sido comparado con un nivel de referencia de  $\chi^2_{(3)} = 7,815$ . Por lo tanto, no hay suficiente evidencia al nivel del 5% para suponer que los datos no provienen de una distribución normal de parámetros 6472,069 y 831,562<sup>2</sup>.

La Figura 5, y luego de realizar el test anterior, nos permite tener una visión gráfica del ajuste, permitiendo comparar el histograma con la campana de Gauss y poder visualizar gráficamente que se aproximan.

A partir de lo expuesto anteriormente, pusimos a prueba nuestra hipótesis mediante técnicas inferenciales, en este caso utilizamos un Test de Hipótesis para la media de una distribución Normal con varianza desconocida.

A continuación, R nos informa del valor del estadístico de contraste ( $t = 1,1206$ ), de los grados de libertad ( $df = 84$ ) y del p-valor ( $p\text{-value} = 0.2657$ ). Ya podemos, por tanto, concluir: dado que el p-valor no es inferior al 5%, no tenemos suficientes evidencias en los datos para rechazar la hipótesis nula ( $\mu = 6371$ ) en favor de la alternativa ( $\mu \neq 6371$ ), es decir, con los datos de la muestra no tenemos suficientes evidencias de que el radio de la Tierra sea distinto de 6371 km.

También, R proporciona un intervalo de confianza del 95%, para la media de la distribución normal que se le supone a los datos, de manera que podemos concluir que el radio de la tierra se estima, con una probabilidad del 95%, entre 6292.705 km y 6651.432 km. Existe una estrecha relación entre este intervalo y el contraste de hipótesis, ya que el valor hipotético que hemos considerado para la media, 6371 km, está dentro de este intervalo.

## Conclusión

Uno de los objetivos de estudio de este trabajo consistió en probar si efectivamente el radio calculado por las instituciones que participaron del Proyecto Eratóstenes 2012 difiere del calculado científicamente. Luego de distintos análisis pudimos llegar a la conclusión de que los datos muestrales ofrecen suficiente evidencia para suponer que la distribución de los datos es Normal y que el valor calculado para la media muestral no difiere significativamente del valor aceptado hoy en día.

En cuanto al otro objetivo, se intentó describir en forma resumida y con los argumentos suficientes, el método utilizado por Eratóstenes al momento de calcular el perímetro y el radio de la Tierra 2200 años atrás.

## Referencias

- Carbajal, M.. (2012). *El proyecto Eratóstenes: La astronomía de los pueblos originarios*. Departamento de Física. UBA. [página web]. Disponible en <http://www.df.uba.ar>. [Agosto, 2015]
- Gómez Aroca, J.M. (20 de noviembre 2015). *Eratóstenes y la radio de la Tierra*. [página web]. Disponible en <http://mimosa.pntic.mec.es/jgomez53/matema/practica/eratostenes.htm>. [Agosto 2015] .
- Sáez Pastor, F. (2010). *Medir el radio de la Tierra. Una experiencia educativa*. Revista de Investigación en Educación, 8, pp.140-146
- R Core Team (2015). R. *A language and software environment for statistical computing*. R Foundation for Statistical Computing, Viena, Austria. [software] Disponible en [www.r-project.org](http://www.r-project.org).

## EFEECTO DE LESIONES CEREBRALES

Daira Bernardez - Julieta Curto

[bernardezdaira@gmail.com](mailto:bernardezdaira@gmail.com) - [curtojulieta@gmail.com](mailto:curtojulieta@gmail.com)

### Resumen

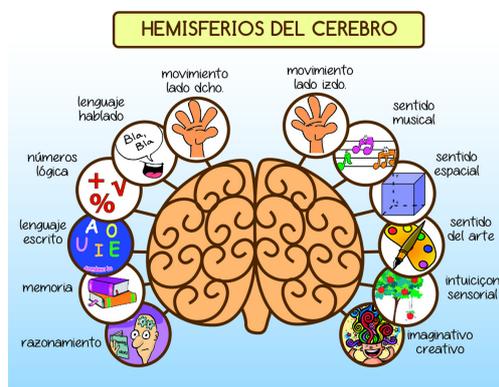
En este trabajo se presenta un análisis sobre los efectos de lesiones unilaterales del hemisferio izquierdo y el hemisferio derecho. Para esto se consideraron tres poblaciones de estudio: pacientes que sufrieron daño cerebral en el hemisferio izquierdo, aquellos con lesiones cerebrales en el hemisferio derecho y los pacientes de control. Se realizó un estudio descriptivo e inferencial para determinar si la exactitud y precisión de la respuesta del paciente se ve afectada por lesiones cerebrales unilaterales, independientemente de su lado.

### Introducción

El cerebro humano está dividido en dos hemisferios cerebrales, el hemisferio derecho y el hemisferio izquierdo. Estos hemisferios están unidos por una especie de puente que llamamos “cuerpo calloso”. Gracias a él estos podrán comunicarse, trabajar en armonía y complementarse.

El hemisferio izquierdo se encarga del habla, la escritura, la numeración, las matemáticas, la lógica, los pensamientos analíticos y el control del lado derecho del cuerpo, como también de las facultades necesarias para transformar un conjunto de información en palabras, gestos y pensamientos. Una lesión en este hemisferio puede producir pérdidas funcionales, pérdida funcional del habla y afectar destrezas motoras en el lado derecho del cuerpo, así como también influir en la exactitud al elegir.

El hemisferio derecho del cerebro, en cambio, controla la comunicación no-verbal, las emociones, y la orientación espacial o la sensación de la posición del cuerpo. Este hemisferio es el centro de las facultades viso-espaciales no verbales, especializado en sensaciones, sentimientos, y habilidades especiales como las visuales y sonoras. Es el responsable de nuestra percepción del mundo en términos de color, forma y lugar. Una lesión en este hemisferio puede causar varios problemas en las emociones y en el comportamiento de la persona, como también en el control del lado izquierdo del cuerpo y en el reconocimiento de objetos, lugares y personas.



**Figura 1:** Funciones de los hemisferios cerebrales.

La hipótesis planteada en este proyecto es que los pacientes con daño cerebral en el hemisferio izquierdo tendrán menos precisión al responder, ya que las personas que sufren estas lesiones pueden ser lentas al actuar o al responder preguntas, comparado con aquellos pacientes con daño cerebral unilateral en el hemisferio derecho. Por ello, el objetivo del presente trabajo es evaluar si la precisión de la respuesta del paciente se ve afectada por lesiones cerebrales unilaterales, independientemente de su lado.

## Metodología

En este trabajo se examinaron los efectos de lesiones cerebrales unilaterales en el hemisferio izquierdo y el hemisferio derecho en la exactitud para elegir la respuesta correcta en preguntas con cuatro opciones cada una de ellas.

Para este estudio se consideraron tres poblaciones: pacientes con daño cerebral en el hemisferio derecho, pacientes con daño cerebral en el hemisferio izquierdo y pacientes de control.

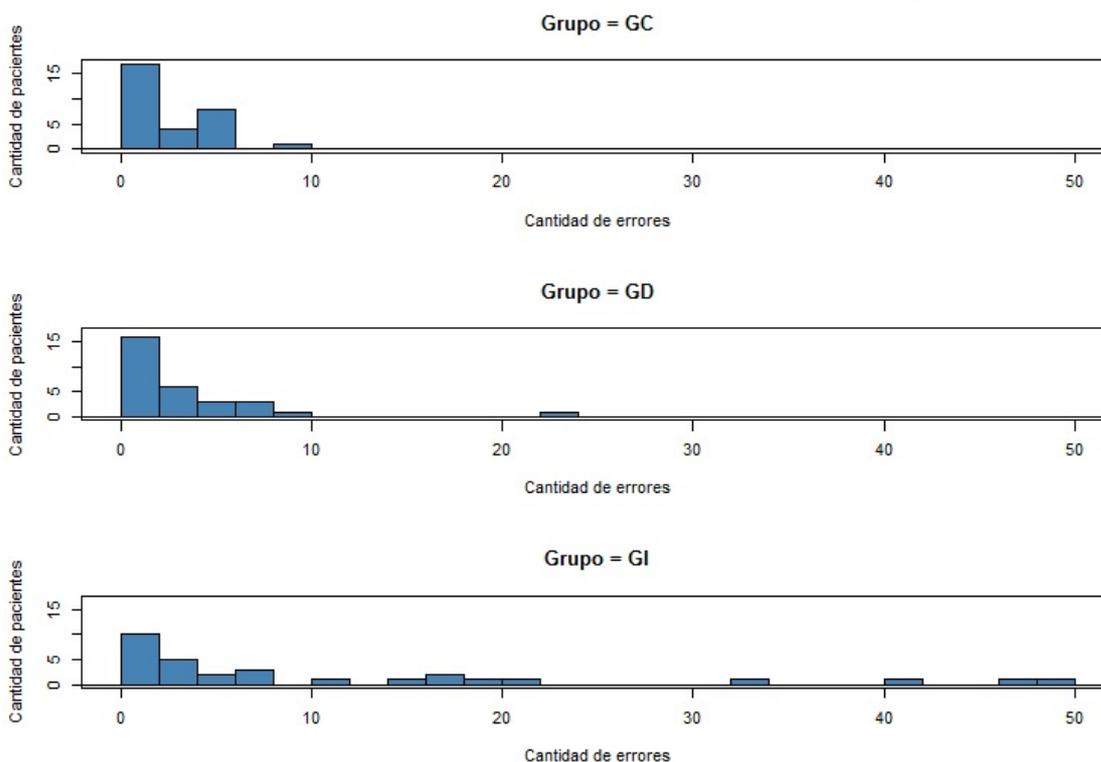
En base a estas poblaciones se seleccionaron tres muestras aleatorias simples independientes, 30 pacientes con lesión cerebral en el hemisferio izquierdo (GI), 30 pacientes con lesión cerebral en el hemisferio derecho (GD), y 30 pacientes de control (GC). En todos los casos se registró la cantidad de errores que cometieron al responder las preguntas que se les formularon en la primera fase del experimento.

Se muestran los resultados obtenidos del análisis descriptivo e inferencial (test de hipótesis de nivel asintótico para la diferencia de medias de las muestras GI y GD) realizado sobre las tres poblaciones a analizar mediante la utilización del software R mediante su interfaz gráfica RCommander.

## Resultados

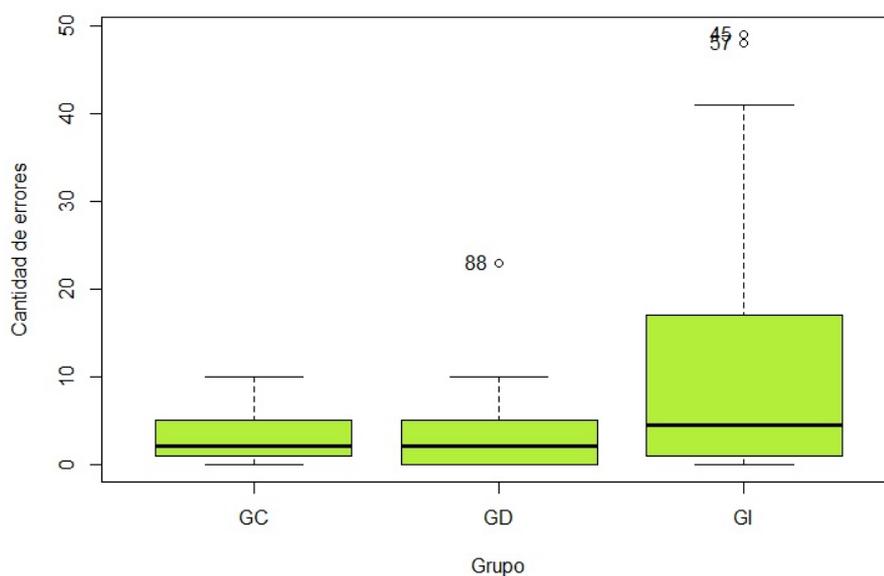
Con los resultados obtenidos respecto de la cantidad de errores producidos por los pacientes en la primera fase del experimento correspondientes a cada grupo se realizaron los cálculos relacionados a la estadística descriptiva, entre ellos, histogramas, diagramas de cajas (Boxplot) y medidas de resumen necesarias para el análisis.

### Histograma de errores producidos por individuos en una fase del experimento



**Figura 2:** Histograma por grupos de la cantidad de errores producidos por los pacientes en la primera fase del experimento.

### Boxplot de errores producidos por individuos en una fase del experimento



**Figura 3:** Boxplot por grupos correspondientes a la cantidad de errores producidos por los pacientes en la primera fase del experimento.

En el Cuadro 1, presentamos las medidas de tendencia central y de dispersión correspondientes a cada muestra analizada.

MEDIDAS	GC	GD	GI
<i>Media</i>	2.93 errores	3.33 errores	10.97 errores
<i>Desvío</i>	2.35 errores	4.69 errores	14.34 errores
<i>Rango intercuartilico</i>	4.00 errores	4.75 errores	15.50 errores
<i>Mínimo</i>	0 errores	0 errores	0 errores
<i>Máximo</i>	10 errores	10 errores	41 errores
<i>Coefficiente de variación (CV)</i>	0.80	1.41	1.31
<i>MAD (mediana de los desvíos absolutos)</i>	1.99 errores	3.24 errores	11.26 errores
<i>Cuantil 25%</i>	1 errores	0 errores	1 errores
<i>Cuantil 50% (mediana)</i>	2.0 errores	2.0 errores	4.5 errores
<i>Cuantil 75%</i>	5 errores	4.75 errores	16.50 errores

**Cuadro 1:** medidas de tendencia central y de dispersión.

Teniendo en cuenta el conjunto de datos y las variables de estudio correspondientes a cada población,

GD = Número de errores que comete un paciente con daño cerebral unilateral derecho,

GI = Número de errores que comete un paciente con daño cerebral unilateral izquierdo,

GC = Número de errores que comete un paciente de control,

podemos determinar su distribución binomial con los siguientes parámetros, siendo  $k$  la cantidad de preguntas a responder por cada paciente en la primer fase del experimento y  $p_i$  con  $i = 1, 2, 3$ , la probabilidad de cometer un error en cada pregunta.

Al realizar los test de hipótesis de nivel asintótico obtuvimos  $p = 1.143802 e^{-14}$  rechazando la hipótesis nula al ser  $p < \alpha = 0.05$ , por lo tanto se construyó el IC correspondiente  $IC (\mu_{GI} - \mu_{GD} : 95\%) = [3.702313 ; 11.564354]$ , el cual indica que la media de la cantidad de errores según un daño cerebral en el hemisferio derecho o en el hemisferio izquierdo difiere, ya que este intervalo no incluye al cero, así como también podemos asegurar que la diferencia de las medias de GI y GD, se encontrarán dentro de estos valores.

## Discusión

En la Figura 2 podemos observar que en el grupo de control (GC) los errores producidos por cada paciente varían entre 0 y 10, al igual que en los pacientes con daño cerebral en el hemisferio derecho (GD), excepto uno de ellos el cual tuvo 23 errores en la primera fase del experimento. En ambos grupos la mayor cantidad de individuos produjo entre 0 y 2 errores.

Observando el histograma, a simple vista podemos determinar que la cantidad de errores en las personas con daño cerebral en el hemisferio izquierdo (GI) tiene mayor variabilidad (0 a 50 errores). El porcentaje más alto de los 30 pacientes que pertenecen al grupo GI generan errores entre 0 y 22, sólo cuatro de estos pacientes producen de 32 a 50 errores.

Los diagramas de caja y brazos o boxplot en paralelo de la Figura 3 permiten comparar la distribución de los errores según la población a la que pertenecen. Observamos que la distribución de la cantidad de errores producidos por el grupo GC y GD es similar en ambos, coincidiendo sus máximos y mínimos. Asimismo dentro del grupo de control podemos observar a un paciente particular (observación 88) con una cantidad de errores atípica, al igual que los dos pacientes con daño cerebral izquierdo los cuales también tienen una cantidad atípica de errores dentro de su grupo (observación 45 y 57).

Podemos determinar que la distribución de la cantidad de errores producidos por los pacientes pertenecientes al grupo GI es altamente asimétrica positiva, registrándose una mayor concentración de datos por debajo de 4.5 errores y una disgregación de valores por sobre esta cantidad.

Siguiendo con la distribución de los errores del grupo GI en la Figura 3, la asimetría de la cantidad de errores obtenidas y analizadas obligan a utilizar como medida de tendencia central asociada la mediana, siendo la MAD (sin corrección) la medida de dispersión asociada. Luego la cantidad media de errores producidas por GC, GD y GI resulta  $(2 \pm 1,99)$ ,  $(2 \pm 3.24)$  y  $(4.5 \pm 11.26)$  respectivamente. Podemos decir que la cantidad media de errores producidos por el grupo GI es más alta y la distribución es más dispersa.

Teniendo en cuenta nuestras variables de estudio determinadas anteriormente:

$GD \sim \text{Bi}(k; p_1)$ , siendo  $GD_1 \dots GD_{30}$  una m.a.s independiente de GD.

$GI \sim \text{Bi}(k; p_2)$ , siendo  $GI_1 \dots GI_{30}$  una m.a.s independiente de GI.

$GC \sim \text{Bi}(k; p_3)$ , siendo  $GC_1 \dots GC_{30}$  una m.a.s independiente de GC.

Para comprobar nuestra hipótesis (los pacientes con daño cerebral en el hemisferio izquierdo tendrán menos precisión al responder), utilizamos el TH con nivel asintótico para diferencia de medias con un tamaño muestral suficientemente grande ( $n=30$ ) al nivel de significación  $\alpha = 0.05$  dado que nuestras variables siguen una distribución binomial y no una distribución normal. Si observamos los histogramas de la Figura 2, es evidente que estos no cumplen con la gráfica esperada de una distribución binomial. Sin embargo, con los datos muestrales de nuestras variables de estudio (Grupo y Cantidad de errores), podemos asegurar que las variables efectivamente siguen dicha distribución, pero con ciertas particularidades ( $n$  grande y  $p$  cercano a cero:  $n > 30$  y  $np \leq 5$ ).

Comprobamos que el grupo GI tiene mayor cantidad de errores que el grupo GD.

Sabemos que:

$$E(GD) = kp_1 = \mu_{GD}$$

$$E(GI) = kp_2 = \mu_{GI}$$

$$E(GC) = kp_3 = \mu_{GC}$$

Tomando como pivote  $T = \frac{(\overline{GI} - \overline{GD}) - (\mu_{GI} - \mu_{GD})}{\sqrt{\frac{s_{GI}^2 + s_{GD}^2}{n}}}$  el cual tiende en distribución a una distribución

normal estándar por el Teorema Central del Límite y el Teorema de Slutsky.

Se determinó analíticamente a través de la construcción del test de hipótesis con nivel asintótico unilateral, que la cantidad de errores generados por los pacientes con daño cerebral izquierdo fue mayor que la cantidad de errores cometidos por los pacientes con daño cerebral derecho. Tomando el pivote  $T$ , trabajando con un nivel de significación del 5% y poniendo a prueba las siguientes hipótesis:

$$H_0: \mu_{GI} - \mu_{GD} \leq 0 \quad \text{vs} \quad H_a: \mu_{GI} - \mu_{GD} > 0$$

Obteniendo un  $p = 1.143802 e^{-14}$ , siendo  $p \leq \alpha = 0.05$  podemos concluir que se rechaza la hipótesis nula ( $H_0$ ) al nivel del 5%. Además, la región de rechazo obtenida es  $RR = \{ \overline{GI} - \overline{GD} > 3.702313 \}$  y siendo  $\overline{GI} - \overline{GD} = 7.64 \in RR$ , luego hay suficiente evidencia para rechazar  $H_0$  i.e. que la cantidad de errores producidos por GI es mayor que la cantidad de errores de GD.

Al haber rechazado la hipótesis nula, se construye el intervalo de confianza (IC) de nivel asintótico para la diferencia de medias con un nivel de significancia del 5%:

$$IC(\mu_{GI} - \mu_{GD} : 95\%) = [3.702313 ; 11.564354]$$

Este intervalo nos permite encontrar valores factibles para nuestro parámetro de interés, pudiendo concluir que la diferencia de números de errores producidos por personas que tienen lesiones en el hemisferio izquierdo y el derecho ronda las 4 y 11 unidades, con una probabilidad del 95%.

Por lo tanto, como el IC construido no contiene al cero, se ratifica nuestra conclusión anterior,

## Conclusiones

A partir del estudio realizado a 90 personas, 30 pertenecientes al grupo de control, 30 con daño cerebral unilateral derecho y 30 con daño cerebral unilateral izquierdo, se concluye mediante la parte descriptiva del análisis y a través de la inferencia estadística (test de hipótesis de nivel asintótico) con un nivel de confianza del 95% que la cantidad de errores de los pacientes que pertenecen al grupo GI es mayor que los errores generados por el grupo GD. En resumen, un daño cerebral unilateral en el hemisferio izquierdo afecta en mayor medida la precisión y exactitud de respuesta en una persona.

## Referencias

- Cabrera Olayo, D. (21 de mayo de 2009). *4 Diferencias entre el hemisferio izquierdo y el derecho*. [página web]. Disponible en: <http://danieljcabrera.com/2009/05/21/4-diferencias-entre-el-hemisferio-izquierdo-y-el-derecho/> [30 de septiembre de 2015]
- Castillo Sáez, A.J. (2009). *Métodos estadísticos con R y R Commander*. Departamento de Estadística e Investigación Operativa. Universidad de Jaén [recurso web].
- Federación Española de Daño Cerebral (2000). *Daño cerebral. Guía de familias*. Madrid: Ministerio de Trabajo y Asuntos Sociales. Secretaría General de Asuntos Sociales. Instituto de Migraciones y Servicios Sociales [libro].
- Joanette, Y.; Ansaldo, A.I.; Abusamra, V.; Ferreres, A.; Roch-Lecours, A.; Kahlaoui, K.; Côté, H. (2008). *Impacto de las lesiones del hemisferio derecho sobre las habilidades lingüísticas: perspectivas teórica y clínica*. Rev Neurol, 46 (8): pp. 481-488.

- Pérez Martínez, DA.; Puente Muñoz, AI; Doménech J; Baztán, JJ; Berbel García, A; Porta, Etessam Pérez, J. *Apraxia unilateral del cierre palpebral tras un ictus isquémico: el papel del hemisferio derecho en la comunicación gestual emocional*. Rev Neurol, 44: pp. 411-414.
- R Core Team (2015). Versión 3.2.2 (2015-08-14). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [software] Disponible desde <https://www.R-project.org/>.
- Wikipedia (11 de octubre de 2015). *Hemisferio cerebral*, [en línea]. Disponible en [https://es.wikipedia.org/wiki/Hemisferio\\_cerebral](https://es.wikipedia.org/wiki/Hemisferio_cerebral) [16 de octubre de 2015].
- Wikipedia (28 de septiembre de 2015). *Cerebro humano*, [en línea]. Disponible en [https://es.wikipedia.org/wiki/Cerebro\\_humano](https://es.wikipedia.org/wiki/Cerebro_humano) [30 de septiembre de 2015].

## LEY DE BENFORD

Carlos Sebastián Cerani Orosco - Manuel Olivera  
[seba.cerani@hotmail.com](mailto:seba.cerani@hotmail.com) - [manuelolivera.mo@gmail.com](mailto:manuelolivera.mo@gmail.com)

### Resumen

La Ley de Benford o Ley del Primer Dígito es un fenómeno matemático que predice o asegura que el primer dígito 1 (uno) aparece con mucha más frecuencia que el resto de los números. Esta ley sólo es aplicable a aquellas listas de números que representan datos reales y que no sean totalmente aleatorios. La Ley es aplicable a datos que estén caracterizados por cierto crecimiento exponencial. Particularmente, en este trabajo nos centramos en el análisis descriptivo e inferencial de la población total en la República Argentina distribuida por departamentos, partidos o comunas, de manera tal que, nuestra intención es verificar si este conjunto de 527 datos cumple con la Ley del Primer Dígito.

### Introducción

En 1881, Simon Newcomb (1835-1909), físico, astrónomo y matemático, se percató de que su libro de tablas de logaritmos estaba más desgastado por las primeras páginas que por las últimas, esto le llevó a concluir que en esos años en los que usó ese libro, trabajó más con números cuyas primeras cifras significativas eran bajas. ¿Podía ser más probable encontrar en la naturaleza un número que empezara por 1 o 2, que por 8 o 9? (Martínez Arranz, I. & Torre Sebastián, L. 2005)

En 1938 y de manera independiente, el físico Frank Benford observó el mismo fenómeno en las tablas de logaritmos y realizó una comprobación empírica sobre un total de 20.229 números agrupados en 20 muestras de gran diversidad: áreas fluviales, constantes y magnitudes físicas y químicas, funciones matemáticas e incluso números de direcciones de personas y tomados de portadas de revistas. A partir de los resultados empíricos Benford postuló una “Ley de los números anómalos” para la probabilidad de ocurrencia del primer dígito. Esta ley logarítmica se conoce como “Ley de Benford”. El hecho de que la primera cifra sea la cifra 1 con mayor frecuencia que las demás, puede ser entendido si tenemos en cuenta que comenzamos a contar desde 1 (1, 2, 3,...) hasta llegar al 9, momento en que cada cifra tiene la misma probabilidad. Pero de 10 a 19 sólo tenemos como primera cifra el 1, y sólo cuando llegamos al 99 todas las cifras tendrán la misma probabilidad de nuevo (Wikipedia, 2015a).

Sin embargo, según Guioteca (2015) no fue hasta 1996 que el matemático Ted Hill diera una demostración formal al postulado.

En pocas palabras, la Ley de Benford, también llamada Ley del Primer Dígito, es una ley fenomenológica basada en la distribución de frecuencias de dígitos iniciales en muchos (pero no todos) conjuntos de la vida real de datos numéricos, como por ejemplo facturas, artículos en revistas, direcciones de calles, precios de acciones, número de habitantes, tasas de mortalidad, longitud de los ríos, constantes matemáticas, números primos, entre otros (Wikipedia, 2015b).

Según Estadística para Todos (2008), la Ley tiene múltiples aplicaciones, entre ellas, la detección de la existencia de fraudes fiscales, irregularidades en auditorías internas, distribuciones de espacio

en disco duro, irregularidades en casos clínicos, modelos demográficos, y, ¿quién sabe en qué más casos sería útil?

El primer dígito significativo  $d$ , con  $d=1,2,\dots,9$ , según la Ley de Benford, tiene una probabilidad de ocurrencia de

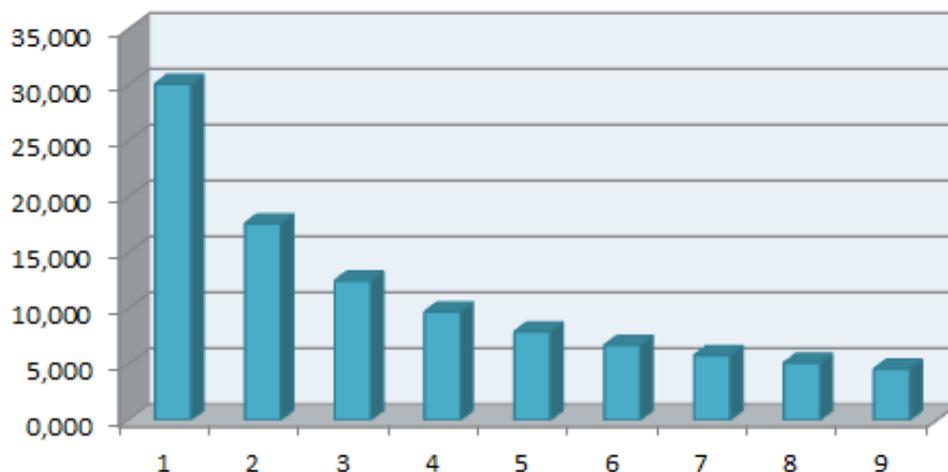
$$P(d) = \log_{10}(d + 1) - \log_{10}(d) = \log_{10} \left( \frac{d + 1}{d} \right) = \log_{10} \left( 1 + \frac{1}{d} \right).$$

De esta manera, se podría construir la siguiente tabla de probabilidades para cada dígito:

Dígito ( $d$ )	Probabilidad ( $P(d)$ )
1	0.301
2	0.176
3	0.125
4	0.097
5	0.079
6	0.067
7	0.058
8	0.051
9	0.046

**Tabla 1:** Probabilidades de cada dígito inicial.

## Distribución de Benford



**Figura 1:** Gráfico de barras de la Distribución de Benford.

En este trabajo se estudia si el conjunto numérico estudiado (Población por departamentos, comuna o partidos) cumple con la Ley de Benford, es decir, si la distribución de frecuencias de los primeros dígitos se asemeja a lo afirmado por la Ley de Benford.

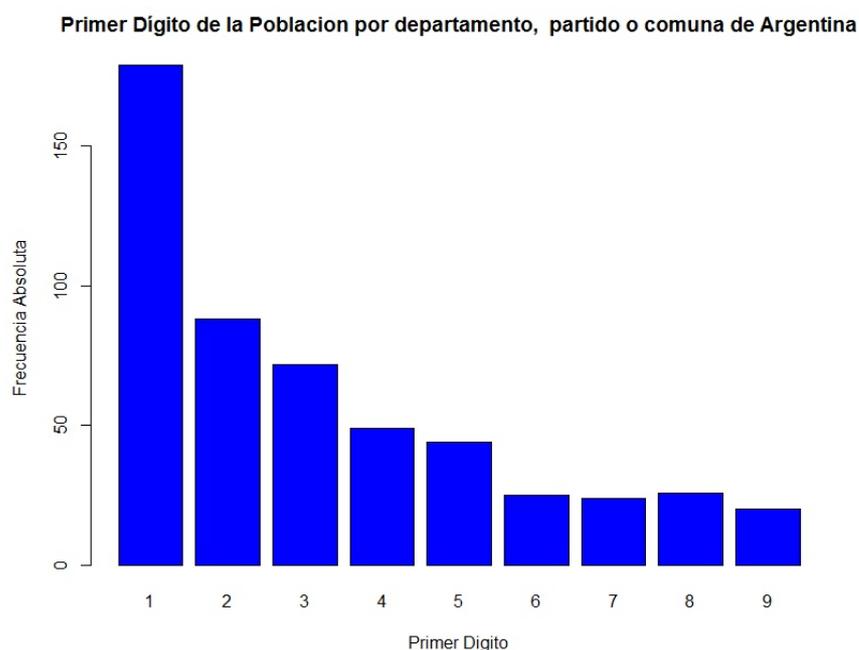
## Metodología

El estudio se realizó sobre la población por departamento, partido o comuna de la República Argentina, la cual la podemos caracterizar como una muestra aleatoria simple de la población por departamento, partido o comuna mundial.

Los datos de las poblaciones por departamento, partido o comuna de la República Argentina, han sido recolectados en la página oficial del Instituto Nacional de Estadística y Censos (INDEC), los cuales surgen del Censo de Población de 2010. En dicho evento, fueron censadas 40.117.096 personas en 527 departamentos, partidos o comunas de la República Argentina. Nosotros consideramos el primer dígito significativo de la población de aquellos departamentos, partidos o comunas. Este primer dígito será nuestra variable a estudiar. Se desarrolló un análisis descriptivo acerca de la frecuencia de cada uno de los dígitos iniciales y se aplicaron técnicas inferenciales. Por el lado descriptivo, el análisis se llevó a cabo a través de gráficos de barras y tablas con medidas de posición y dispersión, y con respecto al análisis inferencial, realizamos test de hipótesis y test de bondad de ajuste. Para ello, utilizamos el software estadístico R 3.2.2 a través de su interfaz gráfica R Commander 2.2-0 y Microsoft Excel 2013.

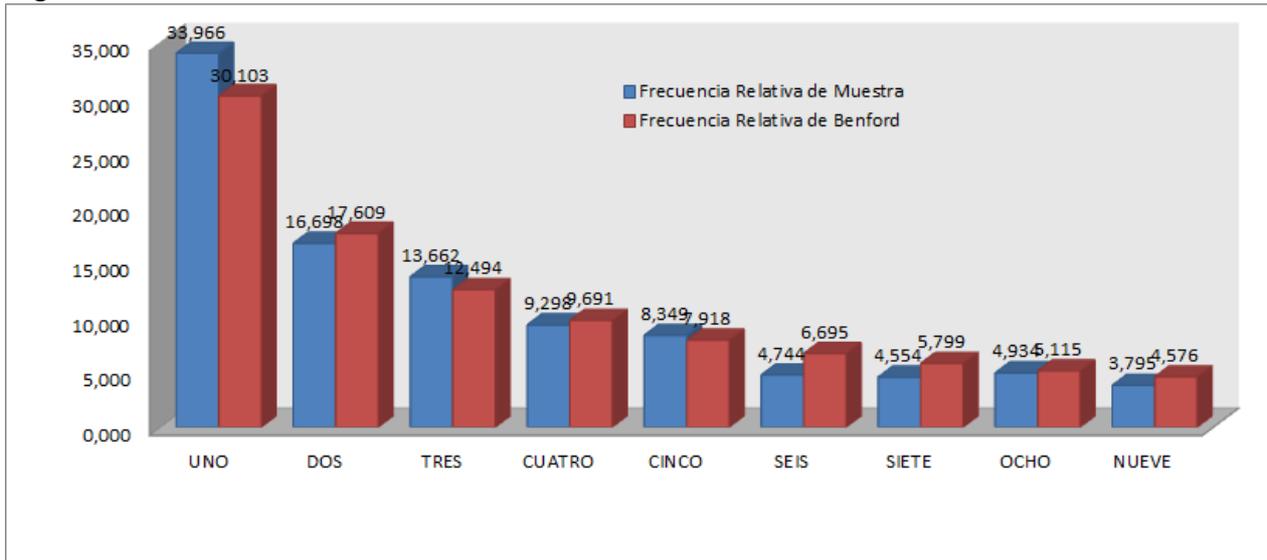
## Resultados

A continuación expondremos un gráfico de barras correspondientes a la frecuencia absoluta del primer dígito de nuestra muestra.



**Figura 2:** Gráfico de barras del primer dígito de la población por departamento, partido o comuna de la República Argentina

Ahora compararemos las frecuencias relativas de la muestra y la distribución de Benford mediante un gráfico de barras elaborado con Microsoft Excel 2013.



**Figura 3:** Comparación entre frecuencias relativas entre la muestra aleatoria simple y la distribución de Benford.

En la siguiente tabla se muestran las distintas medidas de posición y dispersión:

Medidas	Muestra	Benford
Media	3,212524	3,440237
Desvío	2,381886	2,461
Asimetría $\gamma(X)$	0,95114	0,7958
$q_1$	1	1
$q_2$	2	3
$q_3$	5	5
Moda	1	1
Rango Intercuartílico	4	4
Mediana de los desvíos absolutos (MAD)	1,4826	2,9652
MAD Corregido	2,1981	4,396
Coefficiente de variación (CV)	0,741438	0,715358

**Tabla 2:** Medidas de posición y dispersión obtenidas a partir de la muestra y la distribución teórica de Benford.

Analizando los valores encontrados para cada una de las medidas anteriores podemos decir que:

- En ambos casos el primer dígito más significativo que se observa una mayor cantidad de veces es el número uno.
- Interpretando los cuartiles, podemos ver que en los dos casos al menos el 25% de los datos tienen como primer dígito el número uno, y que en el caso de la muestra, el 50% de los primeros dígitos son el 1 y el 2, y en la Ley de Benford, son el 1, 2 y 3. A su vez, ambos acumulan el 75% de los valores hasta el dígito 5. Por esto el RIQ es el mismo en ambas situaciones.
- En ambos casos, se ve que hay una dispersión de los datos muy alta, debido a sus altos coeficientes de variación.
- En ambos casos, la medida más representativa es la mediana, pues las distribuciones, como se ve en la Figura 3, son asimétricas.

Realizamos un test de hipótesis con respecto a la media, con nivel de significación asintótico en base a parámetros estimados por el método de máxima verosimilitud, mediante el cual podemos afirmar que en el 99% de los casos, no hay evidencia suficiente para afirmar que la media teórica y la media muestral son distintas ( $t = -2.194688$ ,  $p = 0.02818599$ ). Para ello, consideramos,  $H_0 : \mu = 3.440237$  vs  $H_a : \mu \neq 3.440237$ , es decir, tomamos como hipótesis nula que la media teórica coincide con la media muestral, y como hipótesis alternativa, que éstas sean distintas. El estimador de la media  $\mu$  es consistente y suficiente.

Usando una prueba de chi-cuadrado de bondad de ajuste, obtuvimos, con un nivel de confianza, es decir, un nivel de significación del 5% los siguientes resultados: valor  $p = 0.3557$ .

## Discusión

A partir de la Figura 2 podemos apreciar que, mientras más pequeño sea el dígito, con mayor frecuencia éste aparece en las poblaciones de los departamentos, partidos o comunas del país. De esta manera, sería lo mismo, pensar que el número 1 aparece más veces que el 2, y el 2 que el 3, y así sucesivamente hasta el 9. Sin embargo, si analizamos, que sucede en el dígito 8, vemos que su frecuencia no es menor que la del 7.

Si analizamos a grandes rasgos, la Figura 1, que proviene de la Distribución de Benford, y la comparamos con la Figura 2, podemos ver que, la única diferencia entre “su forma” es este inconveniente en el dígito 8.

Ahora, si analizamos la Figura 3, podemos ver de manera más precisa, las diferencias entre las frecuencias relativas de la muestra y la Ley de Benford, por dígito. Pero, gráficamente, podemos suponer que las frecuencias de los primeros dígitos en la población por departamento, partido o comuna de Argentina, se aproxima a la Ley de Benford.

En ambas situaciones, la muestreada y la esperada, podemos ver que el 75% de las veces, el primer dígito, está entre 1 y 5. Luego, podemos decir que entre los dígitos iniciales a partir del 5, sólo se acumulan el 25% de los valores. Por otro lado, la mitad de los valores para la muestra se encuentran entre el 1 y el 2, pero, en la Ley de Benford, estos están distribuidos entre los primeros tres dígitos posibles.

Analizando la asimetría a partir de los coeficientes  $\gamma(X)$  encontrados, podemos ver que los valores se encuentran más dispersos hacia la derecha, sólo que en el caso de la muestra, los datos se encuentran aún más dispersos.

Para poder analizar si la media poblacional muestral coincide con la media teórica de la Ley de Benford, se realizó un test de hipótesis con nivel de significación asintótico con respecto a la media. Mediante este test, podemos afirmar que los datos muestrales ofrecen evidencia suficiente para suponer que los datos recolectados tienen una media igual a la de la distribución Benford.

A partir de la prueba de chi-cuadrado de bondad de ajuste realizada, se concluye que la variable de estudio sigue la distribución de Benford, con un nivel de confianza, es decir, un nivel de significación del 5%.

## Conclusiones

El objetivo de este estudio era determinar si se verifica la Ley de Benford en la distribución de la población total de los 527 departamentos, partidos o comunas de la República Argentina. Luego de distintos análisis podemos llegar a la conclusión de que se cumple la Ley de Benford en este conjunto numérico.

## Referencias

- Estadística para todos (2008). *Ley de Benford: Aplicaciones*, [página web]. Disponible en: <http://www.estadisticaparatodos.es/taller/benford/utililidad.html> [11 de noviembre de 2015].
- Guioteca (2013). *Ley de Benford: La increíble técnica matemática para detectar fraudes*, [página web]. Disponible en: <http://www.guioteca.com/matematicas/ley-de-benford-la-increible-tecnica-matematica-para-detectar-fraudes/> [11 de noviembre de 2015].
- Instituto Nacional De Estadísticas y Censos (2012). República Argentina por departamento, partido o comuna. *Población Total. Año 2010*. [página web]. Argentina. Disponible en: <http://www.sig.indec.gov.ar/censo2010/> [20 de octubre de 2015].
- Martínez Arranz, I. & Torre Sebastián, L. (mayo de 2005). *La ley de Benford o el fenómeno del primer dígito*. SIGMA. ISSN 1131-7787, N°. 26, 2005, pp. 131-135. [en línea]. Disponible en: [http://www.hezkuntza.ejgv.euskadi.eus/r43-573/es/contenidos/informacion/dia6\\_sigma/es\\_sigma/adjuntos/sigma\\_26/9\\_ley\\_benford.pdf](http://www.hezkuntza.ejgv.euskadi.eus/r43-573/es/contenidos/informacion/dia6_sigma/es_sigma/adjuntos/sigma_26/9_ley_benford.pdf) [20 de octubre de 2015].
- Microsoft Office 365 (2012). *Microsoft Excel 2013 (15.0.4763.1003)*. Santa Rosa, California [software]. Id.del producto: 00201-10365-53242-AA884.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [software]. Disponible en: <https://www.r-project.org/>.
- Wikipedia (2015a). *Ley de Benford*, [página web]. Disponible en: [https://es.wikipedia.org/wiki/Ley\\_de\\_Benford](https://es.wikipedia.org/wiki/Ley_de_Benford) [20 de octubre de 2015].
- Wikipedia (2015b). *Benford's Law*, [página web]. Disponible en: [https://en.wikipedia.org/wiki/Benford%27s\\_law](https://en.wikipedia.org/wiki/Benford%27s_law) [20 de octubre de 2015].

## CALIDAD DE LOS GENERADORES DE NÚMEROS PSEUDOALEATORIOS

Dasso, Daniela - Mendoza Virgili, Déborah  
[daniela705.1@gmail.com](mailto:daniela705.1@gmail.com) - [mendozavirgili@gmail.com](mailto:mendozavirgili@gmail.com)

### Resumen

Como los seres humanos vivimos en un medio aleatorio y nuestro comportamiento lo es también, para investigar ciertos fenómenos, es necesario tener modelos que simulen la realidad, pero estas herramientas predictivas deben funcionar de manera similar a la naturaleza, es decir, lo más aleatoriamente posible. Es de aquí, que nos surge el interés por determinar la calidad de los generadores de números pseudoaleatorios y establecer cuál es mejor. Para esto, analizamos con argumentos estadísticos, los generadores R, DevC++, GeoGebra y la Calculadora, descriptiva e inferencialmente, tomando una muestra de 100 datos de cada uno. Concluimos que todos los generadores arrojan números pseudoleatorios muy próximos a los aleatorios, destacándose la Calculadora sobre los demás. Como la diferencia de ésta con el resto de los software es mínima, y teniendo en cuenta la rapidez, la eficiencia, la interfaz gráfica y la variedad de herramientas de los software, es más práctico utilizar RCommander o GeoGebra.

### Introducción

Los números aleatorios permiten a los modelos matemáticos representar la realidad.

En general cuando se requiere una impredecibilidad en unos determinados datos, se utilizan números aleatorios. "Los números aleatorios son aquellos que pueden ser generados a partir de fuentes de aleatoriedad, las cuales, generalmente, son de naturaleza física (dados, ruleta, mecanismos electrónicos o mecánicos) y son gobernados por las leyes del azar, éstos exhiben verdadera aleatoriedad en la realización de experimentos. Por su parte los números pseudo-aleatorios son aquellos que tienen un comportamiento similar a la naturaleza aleatoria, pero están ceñidos a un patrón, generalmente de naturaleza matemática, que hace que su comportamiento sea determinístico." (Mancilla Herrera A. M., 2000, p.51)

Como los seres humanos vivimos en un medio aleatorio y nuestro comportamiento lo es también, si deseamos predecir el comportamiento de un material, de un fenómeno climatológico o de un grupo humano podemos inferir a partir de datos estadísticos. Para lograr una mejor aproximación a la realidad, nuestra herramienta predictiva debe funcionar de manera similar a la naturaleza, es decir lo más aleatoriamente posible. De esa necesidad surgieron los modelos de simulación. (Estadística para Todos, 2008).

La simulación es el proceso de diseñar un modelo de un sistema real y llevar a cabo experiencias con él, con la finalidad de aprender el comportamiento del sistema o de evaluar diversas estrategias para el funcionamiento del sistema, comenta Shannon (citado por Tarifa, 2012).

Como ejemplifica Dierker (2015), si quisiéramos determinar si una moneda cualquiera está equilibrada, podríamos lanzarla al aire un número «grande» de veces. Si la frecuencia de caras obtenidas es aproximadamente igual a la de secas podríamos estar suficientemente seguros de que la moneda es legal. Sin embargo, arrojar «muchas» veces la moneda y registrar el resultado podría ser una tarea muy tediosa. Situaciones como ésta, en las que interviene el azar pueden ser simuladas por una computadora.

Dichas simulaciones son las que arrojan números pseudoaleatorios. Los algoritmos que están presentes en computadoras, calculadoras y lenguajes de programación, suelen estar basados en congruencias numéricas obtenidas a partir de ciertos parámetros, por tal motivo permiten generar estos números que parecen aleatorios pero no lo son. Es decir, que son algoritmos determinísticos, a partir de los cuales se obtienen los diferentes resultados bajo las mismas condiciones iniciales, teniendo como denominador común el valor inicial, llamado "semilla". Además, las computadoras tienen una precisión finita, un espacio de memoria determinado, lo cual limita la generación de números con infinita cantidad de dígitos.

Si bien estos métodos para generar números aleatorios nos devuelven otros números que no lo son, su diferencia es indistinguible. Además las ventajas de rapidez, tiempo, sencillez, etc. superan las desventajas antes mencionadas, por eso es que recurrimos a los números pseudoaleatorios en vez de los aleatorios.

“Un **generador de números aleatorios**, es un dispositivo informático o físico diseñados para producir secuencias de números sin un orden aparente” (Wikipedia, 2015b).

Hay diferentes métodos para generar números aleatorios:

- Métodos manuales, como por ejemplo lanzar un dado
- Métodos de Computación analógicas, dependen de procesos físicos – aleatorios
- Métodos de Computación Digital, como por ejemplo R
- Tablas de bibliotecas, números generados por métodos de computación analógica que se han publicado en libros o tablas matemáticas

“Un **generador de números pseudoaleatorios**, es un algoritmo que produce una sucesión de números que es una muy buena aproximación a un conjunto de números aleatorios. La sucesión no es exactamente aleatoria ya que queda completamente determinada por un conjunto relativamente pequeño” (Wikipedia, 2015a).

La mayoría de los algoritmos de generadores de números pseudoaleatorios producen sucesiones que poseen distribución uniforme.

Como señalan Abasolo, Carranza García, De Luna Pérez, Hernández Martínez & Medrano Meza, (2011), los números pseudoaleatorios se generan mediante Algoritmos Determinísticos, divididos en congruenciales y no congruenciales, éstos a su vez se dividen en lineales y no lineales, que requieren parámetros de arranque como una semilla (un valor inicial que se toma como punto de partida para realizar el cálculo del número aleatorio).

Según Wikipedia (2015b), los algoritmos para la generación de valores uniformemente distribuidos están presentes en todas las calculadoras y lenguajes de programación, y suelen estar basados en congruencias numéricas del tipo:

$$x_{n+1} \equiv_{(mod\ m)} (a x_n + c)$$

El éxito de este tipo de generadores de valores de una variable aleatoria depende de la elección de los cuatro parámetros que intervienen inicialmente en la expresión anterior:

- El valor inicial o semilla:  $x_0$
- La constante multiplicativa:  $a$
- La constante aditiva:  $c$
- El número  $m$  respecto al cual se calculan los restos

Estos cuatro valores deben ser números enteros no negativos y que cumplan la siguiente condición:  
 $x_0, a, c < m$ .

Como la mayoría de los algoritmos de generadores de números pseudoaleatorios producen sucesiones que poseen distribución uniforme, lo que nos interesa analizar es la calidad de los algoritmos de **R versión 3.2.1 (2015-06-18)**, la **calculadora CASIO fx-570ES PLUS**, **GeoGebra 5.0.155.0-3D** y **DevC++ versión 4.9.9.2** en base a el ajuste a esta distribución o no. En este sentido, creemos que como el software R es más específico en el ámbito estadístico, nos arrojará números que se ajustan más a una distribución uniforme, es decir “*más aleatorios*” que con el resto de los generadores.

## Metodología

La elección de este proyecto se centra en la importancia de generar números lo más aleatorios posibles. Haremos esta investigación para determinar cuál de los siguientes generadores arrojan números "más aleatorios", o en términos estadísticos cuál de ellos se ajusta a la Distribución Uniforme (0,1). Elegimos los software R, GeoGebra, DevC++, porque son programas con los que trabajamos en distintas materias del Profesorado en Matemática, al igual que la calculadora CASIO fx-570ES PLUS. Los cuales serán ejecutados en una notebook Lenovo modelo B590 con procesador Intel Pentium CPU 2020M, 2.40GHz, disco rígido de 500Gb, memoria Ram de 2Gb y sistema operativo Windows 7 Ultimate de 32 bits.

Recolectamos 100 datos con cada uno de los software y la misma cantidad de números con la calculadora, de la siguiente manera:

- Descargamos el software **R versión 3.2.1 (2015-06-18)** desde <https://www.r-project.org/> e instalamos el paquete **Rcmdr**. Para generar los 100 números pseudoaleatorios en R Commander versión 2.2-0, seguimos la siguiente secuencia *Distribuciones* → *Distribuciones Continuas* → *Distribución Uniforme* → *Muestra de una distribución uniforme...* y luego completamos los campos con 0 el mínimo, 1 el máximo, número de filas 100 y número de columnas 1, damos aceptar y conseguimos el conjunto de datos que queríamos.
- Para generar los números pseudoaleatorios con GeoGebra, primero obtuvimos la versión 5.0.155.0-3D, desde <http://www.geogebra.org/>, luego ingresamos el comando *UniformeAleatorio[ <Mínimo>, <Máximo>, <Número de muestras> ]* en el cuadro de entrada, consiguiendo así la lista de números pseudoaleatorios.
- En DevC++ versión 4.9.9.2, (descargada desde <http://www.bloodshed.net/dev/devcpp.html>) utilizamos el siguiente código:

```
#include <stdlib.h>
#include <stdio.h>#include <time.h>

int main()
{
    int i;
    time_t tiempo = time(0);
    struct tm *tlocal = localtime(&tiempo);
    srand(tlocal->tm_sec+tlocal->tm_min+tlocal->tm_hour+tlocal->tm_year+tlocal->tm_mon+t
local->tm_mday);
    for(i=0; i<100; i++)
    {
        printf("%d\t", rand());
    }
    system("PAUSE");
    return 0;
}
```

Compilamos y ejecutamos para obtener los 100 números pseudoaleatorios. Pero el software nos devuelve números entre el 0 y 32767, entonces primero dividimos esos resultados por 100000 para que nos queden valores entre 0 y 1, y luego le aplicamos una transformación lineal que se corresponda el 0 con el 0 y el 0.32767 con el 1, obteniendo finalmente los datos que buscábamos.

- En la calculadora CASIO fx-570ES PLUS generamos uno a uno los 100 números con las teclas shift → · (Ran#)

Los datos obtenidos serán analizados, comparándolos con la distribución Uniforme (0,1) a partir de gráficos, como histogramas y BoxPlot, medidas de tendencia central y dispersión: la media, el desvío, y parámetros de forma: coeficiente de asimetría y de curtosis, ejecutándolos en el software R y con el paquete RCommander. A partir de tales datos, se realizarán interpretaciones con la finalidad de tomar decisiones o formular predicciones para cumplir con el objetivo propuesto y validar o no la hipótesis.

## Resultados

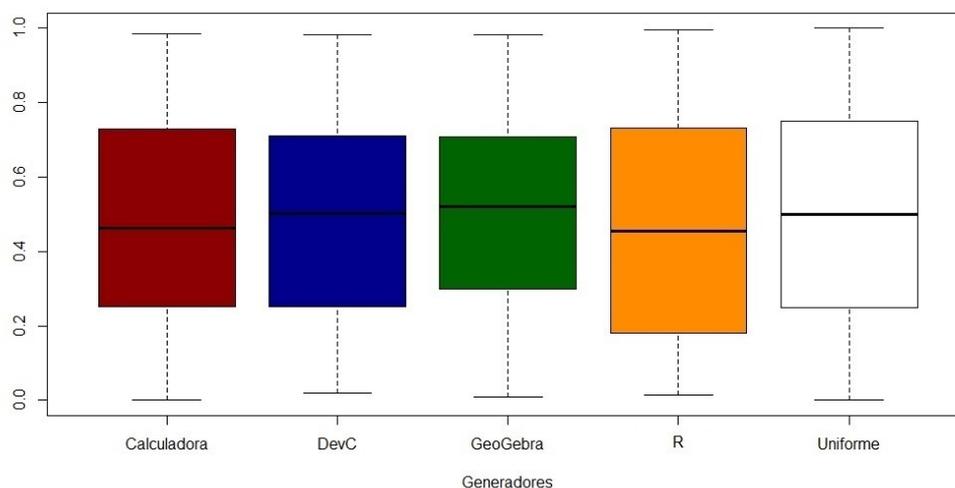
Con las muestras obtenidas, mediante la utilización del software R, a través de su interfaz gráfica RCommander se realizaron varios cálculos, gráficos y test para llevar a cabo la investigación. A continuación presentamos las medidas de tendencia central y dispersión perteneciente a cada uno de los conjuntos de datos obtenidos, para compararlas con los valores teóricos. Pero, que estas medidas coincidan con los valores teóricos de la Distribución Uniforme (0,1) no nos asegura que las variables de los números generados tengan distribución similar a la Uniforme, para ello también necesitamos tener en cuenta los parámetros de forma (curtosis y asimetría).

	Uniforme (0,1)	Datos de R	Datos C++	Datos Calculadora	Datos Geogebra
mean (media)	0.5	0.4730771	0.496885	0.47808	0.5008008
sd (desvío)	0.2886751346	0.2981059	0.2854508	0.2864337	0.267806
Curtosis	1.8	-1.284478	-1.13726	-1.179772	-0.8863847
Asimetría	0	0.09920876	-0.07215526	-0.0108821	-0.01843991
cuartil 0%	0	0.01394744	0.01919614	0	0.01008641
cuartil 25%	0.25	0.1840255	0.2600558	0.25475	0.3026029
cuartil 50%	0.5	0.4559388	0.5030366	0.463	0.520049
cuartil 75%	0.75	0.727682	0.7099445	0.73	0.7066047
cuartil 100%	1.00	0.9966363	0.9814447	0.986	0.9814372
N		100	100	100	100

Figura 1: Medidas de tendencia central y dispersión, parámetros forma de cada una de las variables comparadas con los valores teóricos de la Distribución Uniforme Continua (0,1).

Para analizar con más detalle estas medidas: el máximo, el mínimo y los cuartiles, y poder tener una comparación visual, a continuación se presentarán los Boxplot de cada una de las variables y de la Uniforme (0,1). Los primeros cuatro corresponden a las variables de números generados por la calculadora, DevC++, GeoGebra y R, respectivamente y el último pertenece a una variable con distribución uniforme (0,1).

Figura 2: Comparación de BoxPlot de cada una de las variables con el BoxPlot de la Distribución Uniforme



A continuación, mostraremos los histogramas de densidades para observar la forma que siguen cada una de las variables. De color negro marcamos la función de densidad de una distribución uniforme (0,1), esto nos ayudará a ver hasta dónde deberían ir las barras (correspondientes a los intervalos

determinados por Sturges) de densidad de cada uno de los histogramas.

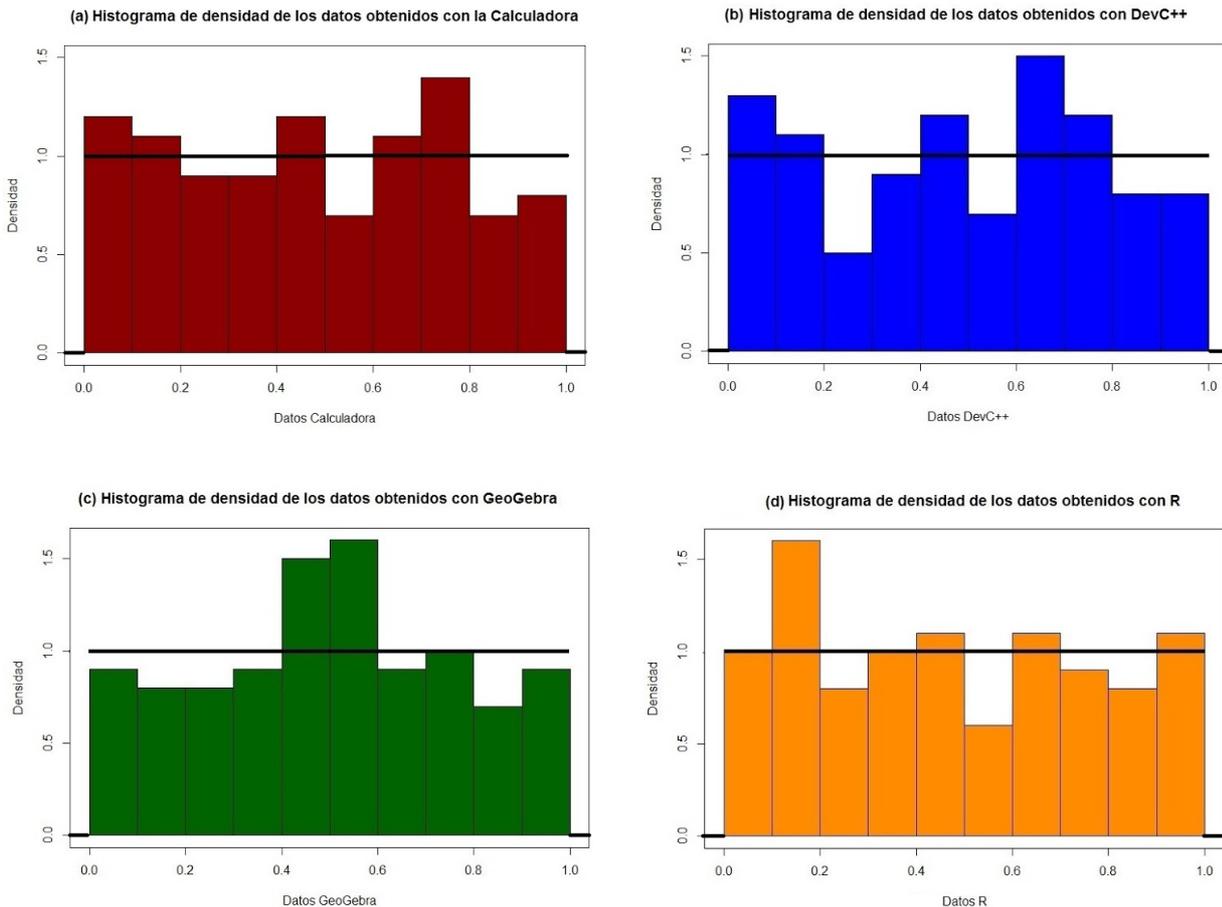


Figura 3: Histogramas de densidad de cada una de las variables y superpuesta la función de densidad de una Distribución Uniforme (0,1)

Para confirmar que las medias de las variables son efectivamente próximas a 0.5 realizamos un Test de Hipótesis para la media con nivel asintótico. Donde determinamos el test bilateral en el cual la hipótesis nula es  $\mu = 0.5$ . Obteniendo los siguientes datos:

- GeoGebra, el estadístico de prueba calculado nos dio  $t = 0.02990052$ , el valor  $p=0.51192681$ , con un coeficiente de confianza del 99%, el intervalo de confianza es de (0.4318185; 0.5697830), lo cual quiere decir que no tengo evidencia para rechazar, por lo tanto aceptamos la hipótesis nula.
- R, el estadístico de prueba calculado nos dio  $t = -0.9031315$ , el valor  $p=0.1832280$  con un coeficiente de confianza del 99%, el intervalo de confianza es de (0.3962901; 0.5498641), lo cual quiere decir que no tengo evidencia para rechazar, por lo tanto aceptamos la hipótesis nula.
- Dev C++, el estadístico de prueba calculado nos dio  $t = -0.1091265$ , el valor  $p=0.4565511$  con un coeficiente de confianza del 99%, el intervalo de confianza es de (0.4233577; 0.5704122), lo cual quiere decir que no tengo evidencia para rechazar, por lo tanto aceptamos la hipótesis nula.
- Calculadora, el estadístico de prueba calculado nos dio  $t = -0.7652732$ , el valor  $p=0.2220544$  con un coeficiente de confianza del 99%, el intervalo de confianza es de (0.4042996; 0.5518604), lo cual quiere decir que no tengo evidencia para rechazar, por lo tanto aceptamos la hipótesis nula.

tanto aceptamos la hipótesis nula.

Como última técnica para definir si las variables siguen Distribución Uniforme Continua (0,1) tal como propusimos, usamos la Prueba Chi cuadrado de Bondad de Ajuste. Consideramos como hipótesis nula justamente que cada una de las variables seguían esa distribución. Y llegamos a la conclusión que todas se ajustan a la distribución supuesta. Particularmente la prueba de bondad sobre los datos obtenidos con el generador R nos dio un valor  $p=0.6993$  y el valor de la chi-cuadrado con 9 grados de libertad resultó 6.4; para la variable de los números generados con GeoGebra el valor  $p$  obtenido fue de 0.5141 y  $\chi^2_{(9)} = 8.2$ ; con la calculadora el valor  $p$  fue de 0.8343 y  $\chi^2_{(9)} = 5$ , por último la Prueba de Bondad de Ajuste para la variable de DevC++ nos arrojó los valores  $p=0.475$  y  $\chi^2_{(9)} = 8.6$ .

## Discusión

Observando la Figura 1, vemos que los números arrojados en general se encuentran entre 0 y 0.9966363 como es de esperarse, además cada variable tiene su media y su desvío aproximado al del modelo uniforme. En este sentido podemos decir que la variable de los números generados con GeoGebra tiene la media más aproximada a 0.5 y el desvío de los números generados con la calculadora es el más aproximado a 0.28867.

También notamos que los parámetros de forma nos dan información bastante parecida a la deseada. Si la distribución es simétrica, el coeficiente de asimetría debe ser cero, y en los casos de nuestras variables este coeficiente es aproximadamente cero. El coeficiente de curtosis debe ser -1.2, para nuestros casos es bastante próximo a ese valor, por lo tanto las variables tenidas en cuenta cuentan con una distribución platicúrtica (acumulan más valores en las colas, que la considerada estándar). Comparando los valores, la variable que tiene el coeficiente de curtosis más próximo a -1.2 es la de los números generados con la calculadora, la cual también tiene el coeficiente de asimetría más cercano a cero.

En cada Boxplot de la Figura 2, podemos observar el máximo, el mínimo y los cuartiles (0.25, 0.5 y 0.75). Los máximos y mínimos de cada Boxplot son muy próximos a 0 y 1, como también lo detallaba la Figura 1. Analizando los cuartiles, tanto desde el punto de vista numérico en la Figura 1, como los Boxplot de cada una de las variables, distinguimos que todas tienden a ser simétricas, acumulando aproximadamente el 25% hasta el primer cuartil, el 50% hasta el segundo y el 75% hasta el tercero, a excepción de la variable de números generados con R que el primer cuartil está por debajo de 0.25 y los generados por GeoGebra que están por encima de 0.25. También podemos observar que la caja más parecida a la uniforme es la de la calculadora, pero ésta no cumple estrictamente con la simetría. La que es más simétrica es la caja de R, pero tiene un brazo más pequeño, lo cual nos habla de la poca dispersión de los datos entre el mínimo y el primer cuartil, y mayor dispersión entre el tercer cuartil y el máximo.

Si bien en un primer momento, observando la Figura 3, pensamos que los histogramas de GeoGebra y R tienen sus barras más próximas a la gráfica de la función de densidad, nos damos cuenta que al tener valores muy extremos, los datos están muy dispersos en comparación al histograma de la calculadora por ejemplo que, aunque no tengan la misma proporción, todos son cercanos a 1.

De los Test de Hipótesis para la media con nivel asintótico que realizamos para cada una de las variables, podemos afirmar que  $\mu = 0.5$ , ya que no tenemos evidencia para rechazar tal hipótesis.

Analizando los Test Chi-cuadrado de Bondad de Ajuste para determinar si las variables se ajustan a una Distribución Uniforme Continua (0,1), podemos concluir que todas las variables siguen una

distribución aproximada a la Uniforme (0,1). Como queremos analizar cuál es mejor, tenemos que tener en cuenta los valores  $p$  que nos arrojó el test en cada una de las variables. En este sentido el  $p$  de la Calculadora es 0.8343, el de R 0.6993, el de GeoGebra 0.5141 y el de DevC++ es de 0.475. Si bien todos se ajustan a la Distribución Uniforme (0,1), lo cual quiere decir que todos generan números pseudoaleatorios pero los números más aleatorios son los generados por la calculadora, ya que su valor  $p$  es mayor que los demás.

## Conclusiones

El objetivo de estudio de este trabajo consistió en establecer cuál de los generadores, R, GeoGebra, DevC++ y la Calculadora, siguen una Distribución Uniforme (0,1), es decir, generan números pseudoaleatorios lo más cercano a números aleatorios.

Nuestra hipótesis es que el software R nos arrojará números que se ajustan más a una Distribución Uniforme (0,1), es decir “*más aleatorios*” que con el resto de los generadores porque es el más específico en el ámbito estadístico.

Luego de distintos análisis, podemos llegar a la conclusión de que los cuatro generadores se ajustan a una Distribución Uniforme (0,1), pero el generador que más se aproxima, es el de la Calculadora, por lo tanto es el de mejor calidad. Si bien pasa esto, la diferencia con el resto de los software es mínima, en consecuencia, por la rapidez, la eficiencia, la interfaz gráfica y la variedad de herramientas, es conveniente utilizar algunos de los software analizados, estableciendo prioridad por el de R, con su paquete RCommander.

## Referencias

- Abasolo, M.A.; Carranza García, A.J.; De Luna Pérez, N.Y.; Hernández Martínez, D.C.; & Medrano Meza, E.R.. (9 de Mayo de 2011). *Números Pseudo-aleatorios y variables aleatorias*. [página web] Disponible en: <http://es.slideshare.net/albertojea/numeros-pseudoaleatorios-y-variables-aleatorias> [23 de Septiembre de 2015]
- Dieser, M.P. (2015). *Probabilidad para futuros matemáticos: Notas de clase*. Material no publicado. Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa. Santa Rosa.
- Estadística Para Todos (2008). *Números Aleatorios*. [página web] Disponible en: <http://www.estadisticaparatodos.es/taller/aleatorios/aleatorios.html> [23 de Septiembre de 2015]
- Mancilla Herrera, A.M. (2000). Números aleatorios. Historia, teoría y aplicaciones. *Ingeniería y Desarrollo*, (8), pp 49-69.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Versión 3.2.1(2015-06-18)) [Software] Disponible en <http://www.Rproject.org/>.
- Tarifa, E.E. (2012). *Teoría de Modelos y Simulación*. [página web] Disponible en: [http://www.econ.unicen.edu.ar/attachments/1051\\_TecnicasIISimulacion.pdf](http://www.econ.unicen.edu.ar/attachments/1051_TecnicasIISimulacion.pdf) [23 de Septiembre de 2015]
- Wikipedia (1 de Marzo de 2015a). *Generador de números pseudoaleatorios*. [página web] Disponible en: [https://es.wikipedia.org/wiki/generador\\_de\\_números\\_pseudoaleatorios](https://es.wikipedia.org/wiki/generador_de_números_pseudoaleatorios) [6 de Octubre de 2015]

Wikipedia (29 de Julio de 2015b). *Generador de números aleatorios*. [página web] Disponible en:  
[https://es.wikipedia.org/wiki/generador\\_de\\_números\\_aleatorios](https://es.wikipedia.org/wiki/generador_de_números_aleatorios) [29 de Septiembre de 2015]

## OBJETOS DEL DISEÑO ERGONÓMICO

Beluzzo, Micaela Belen - Braun, Nerina.

[micabeluzz@gmail.com](mailto:micabeluzz@gmail.com) - [nerinabraun17@gmail.com](mailto:nerinabraun17@gmail.com)

### Resumen

Al utilizar objetos no adecuados a tu antropometría se aumentan los síntomas relacionados a las posturas inadecuadas, movimientos repetitivos y tensiones nerviosas. Por lo cual es importante analizar los objetos del diseño ergonómico de la vida diaria, como lo son los muebles que comúnmente se tienen en una vivienda.

En este trabajo se estudiaron las alturas de la mesa y silla de comedor, y la altura de las camas de las casas de la provincia de La Pampa (L.P.). Las mediciones fueron tomadas en 30 casas de la localidad de Arata (L.P) con el objetivo de verificar si éstas cumplen con las medidas ergonómicas estandarizadas.

### Introducción

Según la *International Ergonomics Association* (IEA, 2001), citado por la Asociación de Ergonomía Argentina (ADEA, 2014) “*La ergonomía es, por un lado, la disciplina científica que busca entender las interacciones entre el hombre y los elementos de un sistema. Por otro lado, es la profesión que aplica en el diseño tanto las teorías, principios, datos, como los métodos para optimizar el bienestar humano y el rendimiento global del sistema. Los ergónomos contribuyen al diseño y la evaluación de tareas, trabajos, productos, entornos y sistemas para que estos sean compatibles con las necesidades, habilidades y limitaciones de las personas*”.

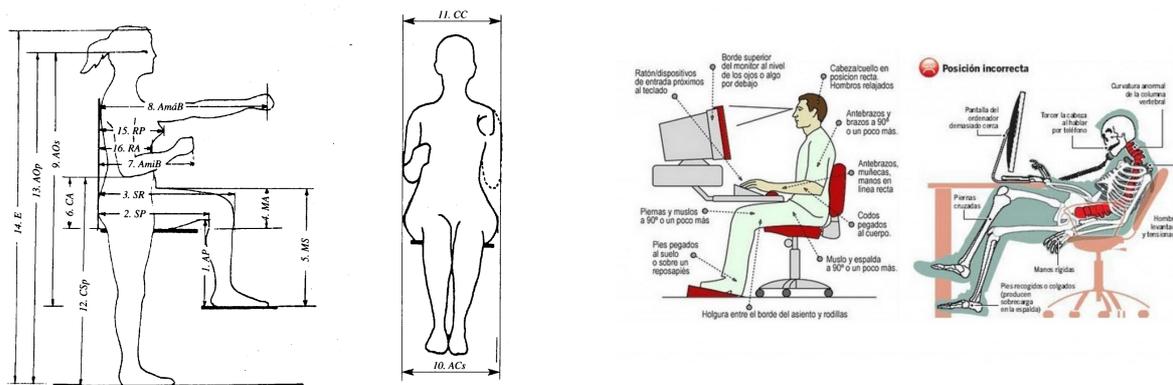
Según Ramis (2009), algunas medidas ergonómicas estandarizadas, son:

- altura mesa comedor: 72 cm.
- altura mesa cocina: 70 cm.
- altura silla comedor: 47 cm.
- altura silla cocina: 45 cm.
- altura mesada: 90 cm.
- espacio en la mesa por persona: 60 cm.
- ancho mínimo de pasillos y escaleras: 90 cm.
- escalones: contrahuella 17 cm y huella 28 cm a 30 cm.
- altura cama: 31 de canapé + 15-16 de colchón : 46-7 cm.

Siguiendo la idea de la Asociación de Ergonomía de la Comunidad Valenciana (s.f.), si tenemos que diseñar un escritorio para un grupo de 10 trabajadores, lo ideal será, como muestra la Figura 1(a), escoger una longitud tal que permita a todos alcanzar el borde final del mismo sin esfuerzo.

Se elige al trabajador que tenga menor alcance de brazos. Si éste pone en riesgo el diseño o la comodidad de los demás trabajadores, se deberá buscar una solución para esta persona (otra mesa adaptada). Lo mismo ocurre para aquél que tenga un mayor alcance de brazos en comparación con los demás trabajadores.

Si queremos diseñar el ancho de una silla, se buscará, como muestra la Figura 1(b) el valor máximo de ancho de caderas para que todos tengan la posibilidad de utilizarla.



a. Medidas que se toman a las personas para fabricar un escritorio y una silla  
by Asociación de Ergonomía de la Comunidad Valenciana

b. Puntos que hay que recordar acerca del diseño del puesto de trabajo by Paula Guarín<sup>4</sup>

Figura 1: Medidas y puntos a tener en cuenta al diseñar escritorio y silla de trabajo.

Como observamos que en nuestra sociedad los síntomas relacionados a las posturas inadecuadas, movimientos repetitivos y tensiones nerviosas son muy frecuentes, decidimos realizar un estudio analizando los objetos del diseño ergonómico relacionados a nuestra vida diaria, como lo son los muebles que comúnmente se tienen en una vivienda.

El objetivo del presente trabajo es verificar si la altura de la mesa y silla de comedor, y la altura de la cama de las casas de la provincia La Pampa (L.P), cumplen con las medidas ergonómicas estándar, pues suponemos que la mayoría de los muebles analizados cumplen con las medidas teóricas.

## Metodología

El estudio se realizó sobre una muestra aleatoria simple, recolectada en 30 de las 460 casas de la localidad de Arata (L.P). Dicha muestra se tomó en esta localidad debido a que conocemos a los propietarios de las viviendas, y al realizar el muestreo, creímos que no surgirían inconvenientes al solicitarles acceder a las mismas.

Como se muestra en la Figura 2, las casas fueron enumeradas por cuadras partiendo de la número 1 ubicada al Nordeste de la localidad, luego se fueron considerando las cuadras recorriendo la localidad de Norte a Sur, al llegar a la última cuadra nos dirigimos a la siguiente ubicada en dirección Oeste, luego recorrimos de Sur a Norte, al llegar a la última cuadra nos dirigimos a la siguiente ubicada en dirección Oeste y continuamos con el mismo criterio. Las casas señaladas con

<sup>4</sup> Guarín, Paula. Portafolio de informática (28 de febrero de 2014). Disponible en: <http://goo.gl/Hx2Dp1> [22 de septiembre de 2015]

un triángulo rojo en el siguiente mapa de la localidad (brindada por la Municipalidad de Arata), corresponden a casas-quinta que fueron consideradas al recolectar los datos.



Figura 2: criterio de numeración de las cuerdas de la localidad de Arata.

La elección del número de la casa para obtener la muestra, se obtuvo a través de la función ALEATORIO.ENTRE [1; 460] del programa Excel 2013<sup>5</sup>.

En cada casa se tomaron las medidas de la altura de la mesa, silla del comedor y cama con una cinta métrica de metal marca Bos de 3 metros.

Como se muestra en la Figura 3, al medir las mesas, se consideró la distancia perpendicular desde el piso hasta el apoyo de la misma, al medir las sillas se consideró la distancia perpendicular desde el piso hasta el asiento, y por último, al medir la altura de la cama se consideró la distancia perpendicular desde el piso hasta el canapé.

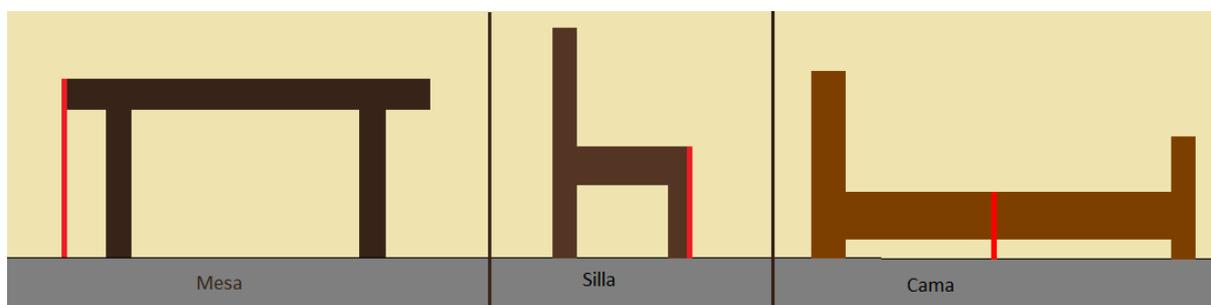


Figura 3: alturas consideradas al medir.

En este estudio no se consideraron las medidas de la mesa y silla de cocina, de la mesada y ancho pasillo, debido a que la mayoría de las casas de la localidad son de barrio, por lo cual no van a diferir las medidas de las mismas. En cuanto al ancho de la escalera y la altura del escalón, también fueron desestimados debido a que en dicha localidad solo hay cinco casas de dos pisos, por lo tanto no fue considerado un dato representativo.

<sup>5</sup> Microsoft Office: Excel (2013)

Se desarrolló un análisis descriptivo sobre cada una de las medidas consideradas y se aplicaron técnicas inferenciales utilizando el software estadístico R Core Team (3.2.2 (2015-08-14)) a través de su interfaz gráfica R Commander (2.2-2)

Para llevar a cabo el objetivo de estudio: establecer si las medidas tomadas de la altura de mesa, silla y cama coinciden con las medidas ergonómicas estandarizadas, primero se realizaron test de hipótesis para la media de cada una de las variables estudiadas con un nivel de 0.05. Además se realizaron intervalos de confianza a un nivel de 0.05 para las medias de cada una de las variables. Por último, para poner a prueba la normalidad de las tres variables estudiadas, se realizó una prueba chi cuadrado de bondad de ajuste con un nivel de 0.01 estimando los dos parámetros.

## Resultados

A continuación expondremos los resultados obtenidos en el análisis de cada una de las variables.

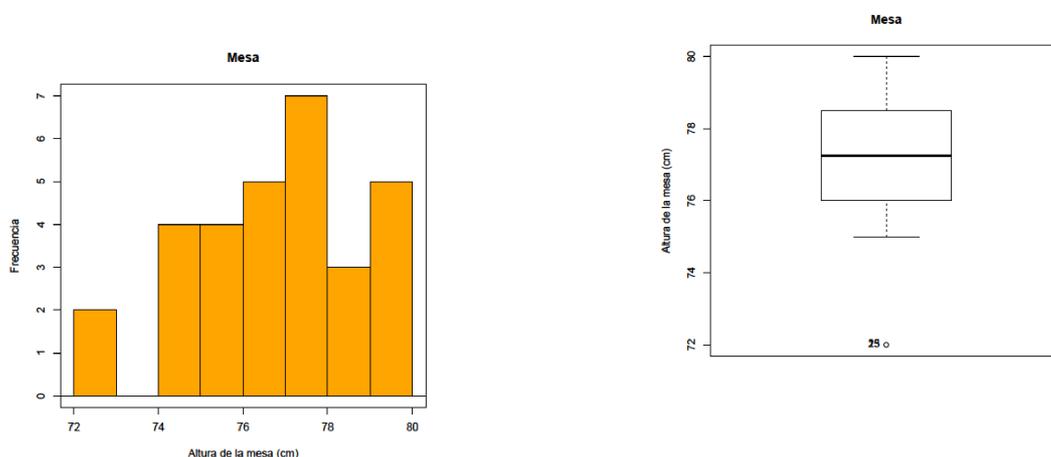
### MESA

Las medidas de tendencia central y de dispersión correspondiente a las medidas de las mesas se sintetizan en la Tabla 1.

Mediana	Media	Desvío	MAD	CV
77.25 cm	77.13333 cm	2.121049 cm	1.85325	0.02749848

Tabla 1: Estadísticos de las alturas de las mesas.

La altura de 30 mesas medidas en centímetros es en media  $77.13333 \pm 2.121049$  cm y en mediana es  $77.25 \pm 1.85325$  cm. Luego como la mediana es más representativa, podemos resumir las alturas de las mesas medidas en centímetros como  $77.25 \pm 1.85325$  cm.



a. Histograma de la altura de las mesas en cm.

b. Boxplot de la altura de las mesas en cm

Figura 4: Gráficos de las alturas de las mesas en 30 casas.

En la Figura 4 (a) podemos observar que la mayoría de las mesas miden entre 77-78 cm, no hay mesas que midan entre 73-74 cm. En cuanto a las que miden entre 74 -75 y 75-76 se observa que tienen la misma frecuencia. La altura máxima tomada fue de 79-80 cm que corresponde a cinco mesas y la mínima fue de 72-73 cm que corresponde a dos.

En la Figura 4 (b) podemos observar una acumulación de datos en los valores más altos, mientras que a los 72 cm se presentan dos puntos extremos que podrían eliminarse, pero se van a considerar ya que no se encontró ninguna anomalía, algo que indique una medición incorrecta o alguna modificación en el mueble. El boxplot no muestra una simetría, sin embargo, si no se consideran los puntos extremos se podría hablar de una distribución simétrica de los datos.

La Figura 4 sugiere que la variable en estudio sigue una distribución normal cuyos parámetros estimados, por el método de máxima verosimilitud, son 77.1333333 y 2.0853990. Los estimadores utilizados son insesgados, eficientes y suficientes. Se realizó un test de bondad chi-cuadrado para poner a prueba la normalidad de la variable: para media y varianza estimadas ( $\chi^2=2.6505$ ,  $gl = 2$ ,  $p= 0.6179$ ).

Para determinar si la altura de las mesas, medidas en centímetros, de las casas de la provincia de La Pampa es igual a 72 cm (medida ergonómica estándar), se realizó un test t para la media al nivel 0.05 ( $p=7.778e^{-14}$ ,  $IC(\mu_m; 0.95)=[76.34132; 77.92535]$ ).

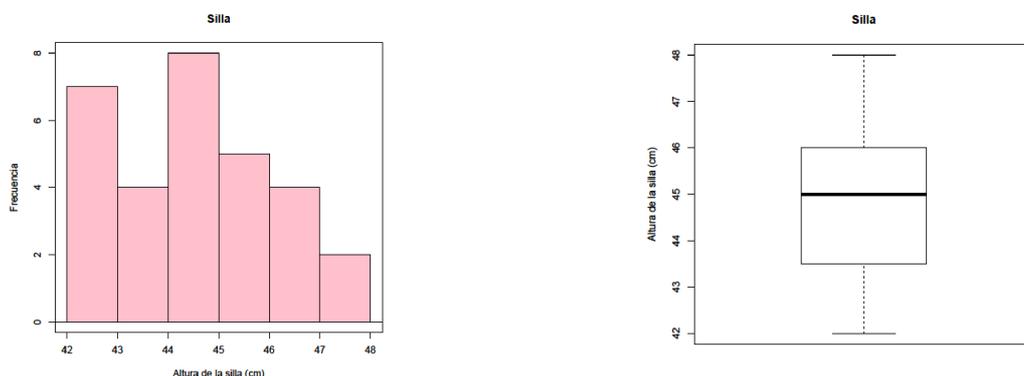
## SILLA

Las medidas de tendencia central y de dispersión correspondiente a las medidas de las sillas se sintetizan en la Tabla 2.

Mediana	Media	Desvío	MAD	CV
45 cm	44.85 cm	1.791503 cm	1.85325	0.03994432

Tabla 2: Estadísticos de las alturas de las sillas.

La altura de 30 sillas medidas en centímetros es en media  $44.85 \pm 1.791503$  cm y en mediana es  $45 \pm 1.85325$  cm. Luego como la media y la mediana no difieren demasiado y la dispersión de los datos no es elevada ( $CV= 3.99\%$ ), es indistinto la elección de una media de posición u otra.



a. Histograma de la altura de las sillas en cm

b. Boxplot de la altura de las sillas en cm

Figura 5: Gráficos de las alturas de las sillas de 30 casas

En la Figura 5 (a) podemos observar que ocho sillas miden entre 44-45 cm, siendo ésta la mayor frecuencia. La altura máxima tomada fue de 47-48 cm que corresponde a dos sillas, siendo esta la menor frecuencia obtenida, y la mínima fue de 42-43 cm que corresponde a siete sillas.

En la Figura 5 (b) podemos observar que hay mayor acumulación en los valores que se encuentran entre la mediana y el tercer cuartil, notando que hay una simetría.

Los gráficos de la Figura 5 sugieren que la variable en estudio sigue una distribución normal cuyos parámetros estimados, por el método de máxima verosimilitud, son 44.85 y 1.7613915. El estimador utilizado es insesgado, eficiente y suficiente. Se realizó un test de bondad chi-cuadrado para poner a prueba la normalidad de la variable para media y varianza estimadas (chi-cuadrado=2.4835,  $gl = 2$ ,  $p = 0.6476$ ).

Para determinar si la altura de las sillas, medidas en centímetros, de las casas de la provincia de La Pampa es igual a 47 cm (medida ergonómica estándar), se realizó un test t para la media al nivel 0.05 ( $p = 3.348e-07$ , IC ( $\mu_s; 0.95$ )=[44.18104; 45.51896]).

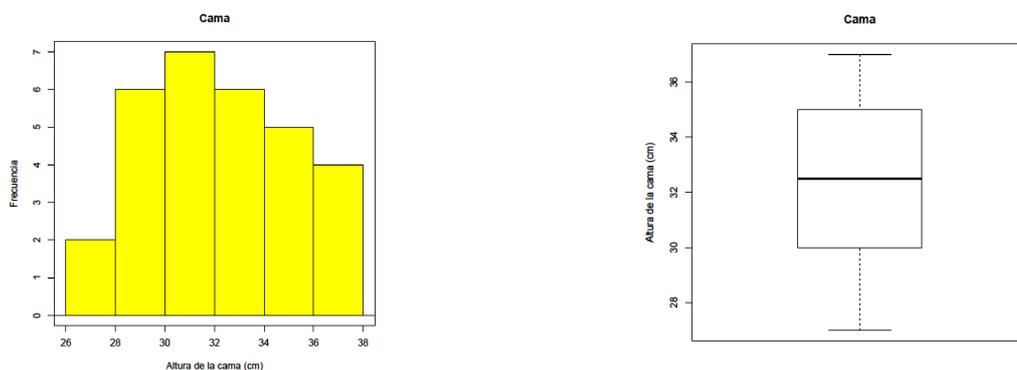
### CAMA

Las medidas de tendencia central y de dispersión correspondiente a las medidas de las camas se sintetizan en la Tabla 3.

Mediana	Media	Desvío	MAD	CV
32.5 cm	32.55 cm	2.98372 cm	3.7065	0.09166575

Tabla 3: Estadísticos de las alturas de las camas.

La altura de 30 camas medidas en centímetros es en media  $32.55 \pm 2.98372$  cm y en mediana es  $32.5 \pm 3.7065$  cm. Luego como la media y la mediana no difieren demasiado y la dispersión de los datos no es elevada ( $CV = 9.17\%$ ), podemos resumir las alturas de las camas medidas en centímetros como  $32.55 \pm 2.98372$  cm.



a. Histograma de la altura de las camas en cm

b. Boxplot de la altura de las camas en cm

Figura 6: Gráficos de las alturas de las camas de 30 casas.

En la Figura 6 (a) podemos observar que siete camas miden entre 30-32 cm, siendo esta la mayor frecuencia. La altura máxima tomada fue de 36-38 cm que corresponde a cuatro camas y la mínima fue de 26-28 cm que corresponde a dos, siendo esta la menor frecuencia obtenida.

En la Figura 6 (b) podemos observar que hay una simetría, ya que la media coincide con la mediana y no se observa una gran dispersión de los datos hacia los extremos.

Los gráficos de la Figura 6 sugieren que la variable en estudio sigue una distribución normal cuyos parámetros estimados, por el método de máxima verosimilitud, son 32.55 y 2.9335701. El estimador utilizado es insesgado, eficiente y suficiente. Se realizó un test de bondad chi-cuadrado para poner a prueba la normalidad de la variable para media y varianza estimadas (chi-cuadrado= 1.6983,  $gl = 2$ ,  $p = 0.791$ ).

Para determinar si la altura de las camas, medidas en centímetros, de las casas de la provincia de La Pampa es igual a 31 cm (medida ergonómica estándar), se realizó un test t para la media al nivel 0.05 ( $p=0.008057$ ; IC ( $\mu_c:0.95$ )=[31.43586; 33.66414]).

## Discusión

Al analizar los datos obtenidos podemos decir que en más del 90% de las medidas observadas, las alturas de las mesas se encuentran entre 74-80 cm, dentro de los cuales se puede observar una mayor acumulación entre 76-78 cm, tal como lo muestran los gráficos de la Figura 4. En el boxplot de la Figura 4 (b) se puede apreciar una dispersión simétrica de los datos con respecto de la mediana, sin considerar los puntos extremos. Por este motivo se supuso que la variable altura de la mesa, medida en centímetros, se distribuye normalmente: el test de bondad de ajuste realizado, revela que hay evidencia al nivel de 1% para suponer que los datos de la muestra provienen de una distribución normal con media 77.13333 y desvío 2.0853990. Luego, podemos resumir a la altura de la mesa, medida en centímetros, como  $77.13333 \pm 2.0853990$  cm.

Algo similar podemos decir sobre las alturas de las sillas. Éstas se encuentran entre 42-48 cm, dentro de los cuales, observando la Figura 5, se puede apreciar una mayor acumulación entre 44-46 cm. A diferencia de lo observado al analizar la altura de las mesas, el boxplot de la Figura 5 (b) no presenta puntos extremos, pudiéndose apreciar una distribución de los datos relativamente simétrica con respecto a la mediana. Por este motivo se supuso que la variable altura de la silla, medida en centímetros, se distribuye normalmente: el test de bondad de ajuste realizado, revela que hay evidencia al nivel de 1% para suponer que los datos de la muestra provienen de una distribución normal con media 44.85 y desvío 1.7613915. Debido a la simetría que presenta la variable, distinguible en la Figura 5, podemos resumir a la altura de la silla, medida en centímetros, como  $44.85 \pm 1.7613915$  cm.

Finalmente, analizando los datos obtenidos podemos decir que las alturas de las camas se encuentran entre 26-38 cm, dentro de los cuales se puede observar una mayor acumulación entre 30-35 cm, tal como lo muestra los gráficos de la Figura 6. Al igual que en el análisis de la altura de las sillas, el boxplot de la Figura 6 (b) refleja que hay una dispersión simétrica de los datos hacia los extremos con respecto de la mediana. Por ello, se supuso que la variable altura de la cama, medida en centímetros, se distribuye normalmente: el test de bondad de ajuste realizado, revela que hay evidencia al nivel de 1% para suponer que los datos de la muestra provienen de una distribución normal con media 32.55 y desvío 2.9335701; pudiendo resumir a la altura de la cama, medida en centímetros, como  $32.55 \pm 2.9335701$  cm.

En cuanto a la relación del estudio de nuestras variables con respecto a las medidas ergonómicas estandarizadas, se realizó un test t de las medias para cada una de dichas variables, por lo que tenemos evidencia suficiente para decir que las medidas de la altura de la mesa, silla y cama de la provincia de La Pampa no cumplen, en media, con las medidas ergonómicas estandarizadas. Además, con un 95% de confianza, podemos afirmar que, en dicha provincia, la altura de la mesa de comedor se encontrarán entre 76.34132 y 77.92535 cm, la altura de la silla de comedor entre 44.18104 y 45.51896 cm, y la altura de la cama entre 31.43586 y 33.66414 cm.

## Conclusiones

El objetivo de estudio de este trabajo consistió en verificar si la altura de la mesa y silla de comedor y la altura de la cama de las casas de la provincia La Pampa, cumplen con las medidas ergonómicas estándar.

Luego de los distintos análisis pudimos llegar a la conclusión de que las medidas ergonómicas estandarizadas no se verifican en dicha provincia. En la mayoría de los casos, las medidas de las mesas de comedor superan la medida estándar encontrándose entre los 76.34 y 77.93 cm; las medidas de las sillas de comedor se encuentran por debajo de la medida estándar, hallándose entre los 44.18 y 45.52 cm; y por último, las medidas de las camas se encuentran entre 31.44 y 33.66 cm, superando los 31 cm establecidos. Sin embargo, en este último caso, es en el que menos difiere la medida ergonómica estándar de la muestral.

## Referencias

- Asociación de Ergonomía Argentina (sin fecha). *¿Qué es la ergonomía?* [página web]. Disponible en: <http://goo.gl/M0hgfl> [22 de septiembre de 2015].
- Asociación de Ergonomía de la Comunidad Valenciana (sin fecha). *Diseño ergonómico y antropometría* [página web]. Disponible en: <http://goo.gl/T21Egb> [22 de septiembre de 2015].
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Software]. Disponible desde: <http://www.R-project.org/>.
- Ramis, M. (septiembre 2009). *La Ergonomía* [página web]. Disponible en: <http://goo.gl/ZeJ6zJ> [05 de noviembre de 2015].