

¡Estadístic@s en acción!

ESTADÍSTICA

PROBABILIDAD Y ESTADÍSTICA II
2016

Patricia GRAGLIA - Evelyn MARSERO - Romina PERALTA

Gastón GALVÁN JUAREZ - Melani MAZZA CALIGARIS

Dahiana GIMÉNEZ - Jorgelina BLANCO - Laura CAFFARONE

Ángel MARTÍN - Eduardo AGÜERO

Mariela WEIS - Cielo URQUIZA

Dalma CAÑADA - Yohana ESNAL - Soledad FELIPE

Compiladoras

Maria Paula DIESER - Lorena Veronica CAVERO



FACULTAD DE CIENCIAS
EXACTAS Y NATURALES

Universidad Nacional de La Pampa

Índice

Prólogo	2
Los tratamientos de la anorexia nerviosa <i>Patricia Graglia, Evelyn Marsero y Romina Peralta</i>	4
La influencia del tabaco en el peso de los recién nacidos <i>Gastón Galván Juárez y Melani Mazza Caligaris</i>	13
Análisis de la relación entre el nivel de colesterol y el infarto de miocardio <i>Dahiana Gimenez, Jorgelina Blanco y Laura Caffarone</i>	19
Relación entre depresión y niveles de ACTH y cortisol <i>Angel Gabriel Martin y Eduardo Daniel Agüero</i>	29
Comparación de proporción de fumadores pasivos y personas no expuestas al humo del tabaco según sus niveles de cotinina <i>Mariela Weis y Cielo Urquiza</i>	37
Hidrocefalia y circunferencia de la cabeza <i>Dalma Cañada, Yohana Esnal y Soledad Felipe</i>	43

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 2016, quienes abordaron diversas temáticas vinculadas con la investigación médica: cuidados prenatales, patologías infantiles, trastornos alimenticios, consecuencias del humo de tabaco en personas no fumadores, entre otras.

Las profes, noviembre de 2016

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/
- 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.

LOS TRATAMIENTOS DE LA ANOREXIA NERVIOSA

Graglia, Patricia - Marsero, Evelyn - Peralta, Romina

patriciamgraglia@gmail.com - evemarsero@gmail.com - romi.s.peralta@gmail.com

Resumen

La Anorexia Nerviosa es un trastorno alimentario que afecta a varias personas, en especial a mujeres adolescentes. Existen diversos tratamientos, entre ellos el cognitivo conductual y el familiar. En este trabajo se presenta un análisis del peso al iniciar y al finalizar los tratamientos, en 72 mujeres jóvenes que padecían este trastorno alimentario. Con esta muestra, se realizó un estudio descriptivo e inferencial para determinar si con estos tratamientos las pacientes subieron de peso y en caso de ser así, con cuál pudieron aumentar más. Luego de compararlos, concluimos que el más apropiado para esto es el familiar y que las pacientes no aumentan de peso si no reciben ningún tipo de tratamiento.

Introducción

En este trabajo se realiza un análisis estadístico sobre los tratamientos a la Anorexia Nerviosa (AN). Esta grave enfermedad tiene múltiples causas, consecuencias, así como tratamientos. Las causas de este trastorno alimentario son un conjunto de factores tanto biológicos, psicológicos como sociales, que afectan principalmente a mujeres, y suele aparecer en la adolescencia, aunque también se está diagnosticando en niñas, pre adolescentes y ancianas. Las personas que padecen esta enfermedad suelen también tener depresión, trastornos obsesivos, y generalmente, su entorno familiar valora mucho la apariencia y el éxito (Borrego, 2000).

Se decidió estudiar este tema porque al trabajar como docentes de nivel secundario, estamos permanentemente en contacto con adolescentes. Sabemos que en esa etapa tan crítica de una persona, se puede ser muy frágil frente a las exigencias de los estereotipos impuestos por la sociedad. Si en algún momento nos encontramos frente a algún caso en el aula, estaremos más informadas sobre el tema. También se puede trabajar de manera interdisciplinaria y, desde el estudio estadístico, se obtiene información relevante para analizar, ayudar a reflexionar y socializar. En 2008, en Argentina, se sancionó la Ley de Trastornos Alimentarios (Ley N° 26396) la cual indica que se debe garantizar asistencia médica, nutricional, clínica y psicológica, además de prácticas quirúrgicas, para quienes padecen alguno de los trastornos alimentarios (obesidad, bulimia, anorexia nerviosa entre otras enfermedades relacionadas). También afirma que el Gobierno debe tomar medidas necesarias para que los anuncios publicitarios y los diseñadores de moda no utilicen a la extrema delgadez como un símbolo de salud, bienestar o belleza. En los kioscos de las escuelas se deben ofrecer y exhibir productos que integren una alimentación equilibrada y variada. Pero, se puede observar que en la actualidad no siempre se cumple con estas disposiciones, por ejemplo, en la mayoría de las instituciones educativas no se ofrecen productos nutritivos.

Al tener tantos factores que originan el cuadro, es de esperarse que también tenga múltiples tratamientos. Entre ellos, se pueden encontrar: tratamientos psicológicos, físicos y farmacológicos.

Al tratar de determinar cuál de los métodos es mejor, se encontraron respuestas poco satisfactorias. El principal problema para determinar la efectividad de uno u otro, es que, por lo general, se utiliza alguna combinación de ellos. El otro inconveniente, según Saldaña (2001), es que la ausencia de grupos de ayuda (para controlar la enfermedad) justificada por la gravedad del trastorno y esta

combinación de tratamientos, han potenciado la escasez de estudios que expliquen las características de las intervenciones más eficaces para el tratamiento de la AN.

En este trabajo, se aborda el análisis de los efectos de dos tratamientos psicológicos: Tratamiento Cognitivo (CBT) y el Tratamiento Familiar (FT); y un tercer caso, en el que las pacientes no reciben tratamiento (Cont). El CBT se trata de que la paciente cambie de pensamientos y actitudes relacionadas con la preocupación por la comida, el peso y la silueta y que luego mantenga los hábitos alimentarios adecuados. Por su parte, el FT se basa en tratar de modificar las relaciones familiares para que la acompañen en todo el tratamiento y la ayuden a organizarse con sus tiempos, sus salidas, formas de alimentarse y de actuar (Morandé, n.d.).

En este trabajo, a estos tratamientos se los analiza por separado. Por ende, según la doctora Matzkin, “no representa la realidad de las pacientes, ya que éstas pueden recibir estos tratamientos pero combinados con uno físico y farmacológico, que de hecho, no necesariamente serán los mismos en todas las pacientes, ya que cada caso es diferente” (V. Matzkin, comunicación personal, 11 de noviembre de 2016). Además, el conjunto de datos tiene solamente el tipo de tratamiento, y los pesos al iniciarlo y al concluirlo, dejando de lado otros factores como el IMC (índice de masa corporal), el peso “ideal” según este índice, la edad al iniciar la enfermedad, etc, que podrían brindar mejor información al pensar cuál tratamiento brinda resultados más satisfactorios.

Por lo tanto, ¿cómo afirmar si un tratamiento es más efectivo que otro? O más aún, ¿cómo determinar si un tratamiento en sí es efectivo? El término “efectivo” es muy amplio y relativo. Por ejemplo, una paciente que al finalizar un tratamiento no aumentó demasiado de peso pero sí ha evolucionado psicológicamente, dicho tratamiento podría considerarse efectivo. Hemos decidido, entonces, no hablar de efectividad de un tratamiento sino hablar de cuánto peso ha aumentado una paciente.

Acorde a los datos disponibles, se realizó un análisis estadístico para determinar si las pacientes aumentan de peso con el CBT, con el FT y sin realizar ningún tratamiento (Cont.). De los tratamientos con los que se aumenta, se quiere determinar además, con cuál de ellos las pacientes aumentaron más de peso. Morandé (n.d.) considera al FT como un “tratamiento a elección”, por lo que nos hace suponer que, si las pacientes se tratan sólo con el FT, no aumentarán de peso o aumentarán menos que con el CBT. Por otro lado, basándonos en la información leída sobre la gravedad de esta enfermedad, tenemos como otra hipótesis que las pacientes sin tratamiento no aumentarán de peso. Saber si con estos tratamientos se aumenta o no de peso es muy importante a la hora de tratar la enfermedad para atender a los objetivos de la paciente. Más aún, si se llega a comprobar que sin tratamiento no se aumenta de peso, sería importante comunicar este trabajo para concientizar a las pacientes de que ésta grave enfermedad requiere de algún tipo de tratamiento.

Metodología

El estudio se realizó sobre una muestra aleatoria simple (m.a.s.), editada por Hand, McConway, Lunn y Ostrowski (1993) en el libro *A Handbook of Small Data Sets*, (p. 229).

El conjunto de datos originalmente estaba formado por una variable cualitativa nominal Tratamiento (CBT - FT - CONT) y dos variables cuantitativas continuas: Previo (peso en libras previo al tratamiento) y Posterior (peso en libras al finalizar el tratamiento). Luego, se agregó una nueva variable cuantitativa, también continua, llamada Diferencia, que representa la resta entre las variables Posterior y Previo.

Se debe remarcar que la unidad de medida de la variable peso es libra y no kilogramos, para no alterar los datos originales y no generar más tipos de errores. Igualmente se consideró que $1\text{ lb} = 0.453592\text{ kg}$.

Para empezar a analizar los datos se realizó, con el software estadístico R (R i386, 3.2.2) mediante su interfaz gráfica RCommander (Rcmdr, 2.3-0), un análisis descriptivo, comenzando por la realización de una tabla de distribución de frecuencias de la variable Tratamiento para saber cuántos datos hay de cada uno de ellos. También se hallaron las medias, los desvíos estándares y los coeficientes de variación de las variables Previo, Posterior y de la nueva variable, Diferencia (creada para poder analizar con cuál tratamiento ganarían más peso) agrupadas por la variable Tratamiento. Según Devore (2005), en diversas investigaciones biométricas se afirma que el peso, la estatura y otras características físicas de las personas se pueden ajustar a una distribución normal. Para analizar la normalidad de las variables Previo y Posterior se confeccionaron los respectivos histogramas y los QQ-plot correspondientes. Como la variable Diferencia es la resta entre dos variables aleatorias dependientes que se ajustan a una distribución normal, entonces la resta entre ambas también se ajustará a una distribución normal tal que $Diferencia \sim N(\mu_1 - \mu_2; \sigma^2)$ donde μ_1 es la media de los pesos previos por tratamiento y μ_2 es la media de los pesos posteriores, $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$ y ρ es el coeficiente de correlación lineal.

A las variables que se ajustan a una distribución normal, se les estimaron los parámetros con el método de máxima verosimilitud. Para realizar estos cálculos y los QQ-plot con la ayuda del software, se crearon tres conjuntos de datos para separar los pesos previos y los posteriores con cada tratamiento.

Para saber si las pacientes aumentaron de peso con cada tratamiento y sin tratamiento, se realizó un test de hipótesis para la diferencia de medias de poblaciones normales para muestras aleatorias dependientes (Previo y Posterior). Y para determinar con cuál de los tratamientos las pacientes subieron más de peso se realizó un test de hipótesis similar con la variable Diferencia según CBT y según FT pero para muestras aleatorias independientes con varianzas desconocidas. Para ello, se realizaron algunos cálculos manualmente y otros con la ayuda del software, creando un nuevo conjunto de datos que contienen sólo los datos para los tratamientos CBT y FT.

Resultados

Análisis descriptivo

Se comenzó confeccionando una tabla de distribución de frecuencias de la variable Tratamiento y se observó que los datos se corresponden con 29 pacientes que se trataron con el CBT, 26 con el FT y 17 que no hicieron ningún tratamiento.

Por otro lado, se calcularon las medidas de tendencia central y de dispersión correspondientes a las variables Previo y Posterior según el tratamiento (ver Tabla 1).

Tratamiento	Previo			Posterior		
	media	s	CV	media	s	CV
CBT	82.69	4.85	0.06	85.70	8.35	0.09
Cont	81.56	5.71	0.07	81.11	4.74	0.06
FT	83.23	5.02	0.06	90.49	8.48	0.09

Tabla 1. Estadísticos de las variables Previo y Posterior por tratamiento.

Como se puede observar en ambas variables, dado que el coeficiente de variación es pequeño,

indica que hay poca dispersión en los datos. Si comparamos las medias de los pesos previos con cada uno de los tratamientos, son muy similares. Sin embargo, las medias de los pesos posteriores ya no son tan similares. Lo que estaría indicando que con algunos tratamientos se sube más de peso que con otros.

Para verificar que la distribución de ambas variables es normal, se realizaron histogramas y los respectivos QQ-plot (ver Figura 1). Como en todos los casos la mayoría de los puntos se acumulan alrededor de una recta, entonces efectivamente provienen de una población normal.

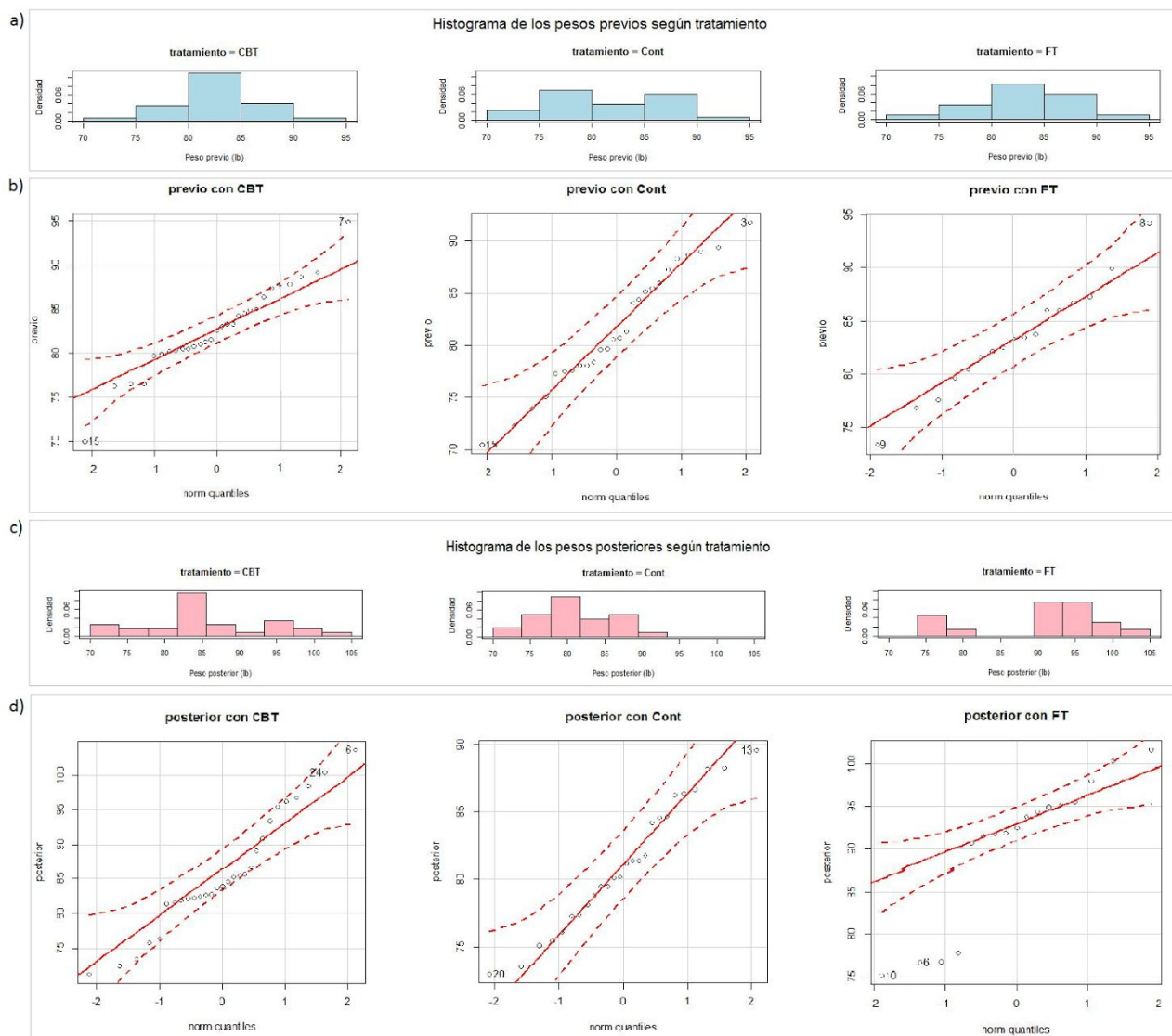


Figura 1. a) Histogramas de los pesos previos por tratamiento. b) QQ-plot de los pesos previos por tratamiento. c) Histogramas de los pesos posteriores por tratamiento. d) QQ-plot de los pesos posteriores por tratamiento.

En cuanto a la variable Diferencia, se obtuvieron la media, el desvío estándar y el coeficiente de variación por tratamiento (ver Tabla 2) y luego se confeccionaron los respectivos histogramas (ver Figura 2).



Tratamiento	Diferencia		
	media	s	CV
CBT	3.01	7.31	2.42
Cont	-0.45	7.99	17.75
FT	7.26	7.16	0.98

Tabla 2. Estadísticos de la variable Diferencia por Tratamiento.

Histogramas para la diferencia de pesos según el tratamiento

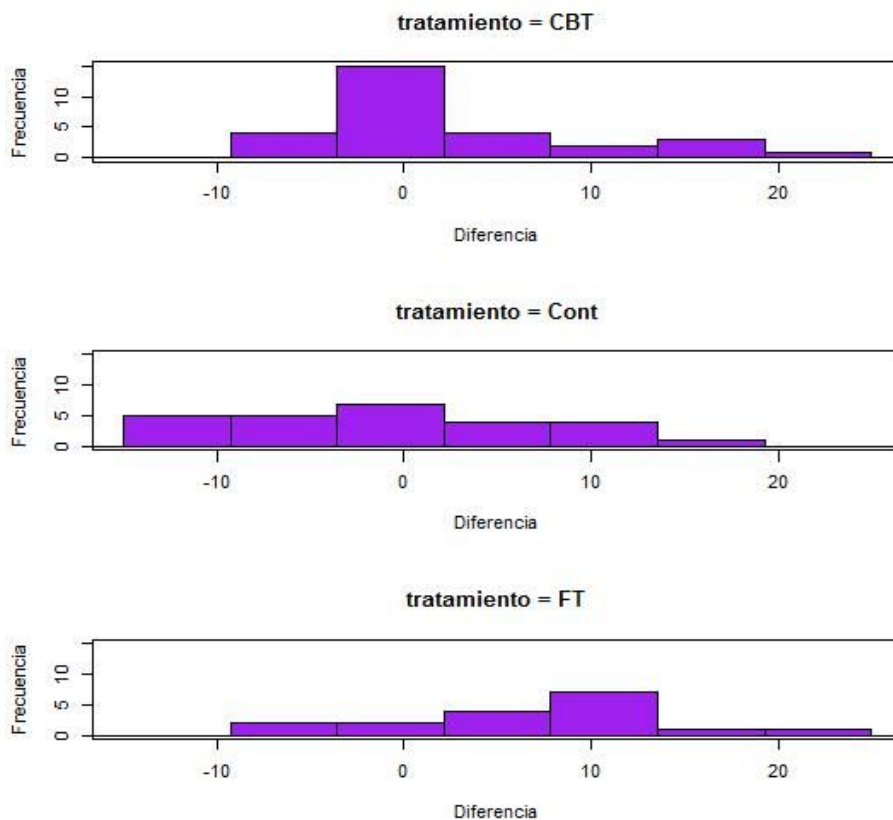


Figura 2. Histogramas de la diferencia de peso según el tratamiento para las 72 pacientes.

Observando los histogramas, las medias con sus desvíos y sus respectivos coeficientes de variación, aunque los datos sean muy dispersos, se podría decir que las pacientes sin tratamiento son las que menos peso ganan y que con el FT han aumentado más de peso.

Análisis Inferencial

Se han estimado los parámetros de la distribución normal con el método de máxima verosimilitud:

$\hat{\mu} = \bar{X}$ y $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ siendo $X_1; X_2; \dots; X_n$ una m.a.s. El estimador $\hat{\mu} = \bar{X}$ es insesgado, eficiente y suficiente (IMVU), mientras que el estimador $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ es asintóticamente

insesgado.

Con ayuda del software se obtuvieron las estimaciones para las medias y las varianzas de las variables Previo y Posterior por Tratamiento. Por lo tanto:

$$\begin{aligned} \text{Previo}_{CBT} &\sim N(82.69; 4.76^2) & \text{Posterior}_{CBT} &\sim N(85.70; 8.20^2) \\ \text{Previo}_{Cont} &\sim N(81.56; 5.60^2) & \text{Posterior}_{Cont} &\sim N(81.11; 4.65^2) \\ \text{Previo}_{FT} &\sim N(83.23; 4.87^2) & \text{Posterior}_{FT} &\sim N(90.49; 8.22^2) \end{aligned}$$

Para determinar si las pacientes aumentaron de peso con cada tratamiento, se realizaron los siguientes test de hipótesis.

➤ Control

Sea D = diferencia entre los pesos (en libras) de las pacientes al finalizar y al comenzar el estudio. Como esta variable es una diferencia de variables que se ajustan a una distribución normal, entonces también se ajustará a una distribución normal $D \sim N(\mu_1 - \mu_2; \sigma^2)$. Por lo tanto el pivote adecuado para realizar el test de hipótesis es $T = \frac{\bar{D} - (\mu_1 - \mu_2)}{\frac{s_D}{\sqrt{n}}} \sim t(n-1)$. Donde s_D es el desvío de la m.a.s.

$D_1; D_2; \dots; D_n$ y n es el tamaño de la muestra.

Se pretende saber si con este tratamiento las pacientes suben de peso, entonces:

$$H_0 : \mu_1 - \mu_2 \leq 0 \quad \text{vs} \quad H_a : \mu_1 - \mu_2 > 0$$

Donde μ_1 es la media de los pesos al finalizar el control y μ_2 es la media de los pesos al comenzar el control. Fijando $\alpha = 0.05$ y habiendo calculado, cuando se realizó el análisis descriptivo, la media muestral y el desvío estándar de la variable *diferencia* para este tratamiento: -0.45 lb y 7.99 lb;

realizando los cálculos manualmente tenemos que la $RR = \{\bar{D} > 2.667\}$ y como $\bar{d} = -0.45 \notin RR$ no hay evidencia suficiente para rechazar H_0 lo que implica que la media del peso posterior no es superior al de la media del peso previo. Por otro lado, analizando el valor p obtenido con el software tenemos que valor $p = 0.61$ y como $\alpha = 0.05 < p$ entonces llegamos a la misma conclusión, no hay evidencia para rechazar H_0 .

➤ Tratamiento Cognitivo Conductual

Sea D = diferencia entre los pesos (en libras) de las pacientes al finalizar y al comenzar el CBT. Realizando un test similar al anterior para saber si con este tratamiento las pacientes aumentan de peso, se tiene que:

$$H_0 : \mu_1 - \mu_2 \leq 0 \quad \text{vs} \quad H_a : \mu_1 - \mu_2 > 0$$

Donde μ_1 es la media de los pesos al finalizar el Tratamiento Cognitivo y μ_2 es la media de los pesos al comenzar con este tratamiento. Tomando el mismo estadístico T , con los valores correspondientes a esta variable y donde s_D es el desvío de la m.a.s. $D_1; D_2; \dots; D_n$ y n es el tamaño de la muestra, fijando $\alpha = 0.05$, realizando los cálculos manualmente, tenemos que la $RR = \{\bar{D} > 2.306\}$. Como $\bar{d} = 3 \in RR$ hay evidencia para rechazar H_0 , lo que indica que la media de los pesos al finalizar el Tratamiento Cognitivo es superior a la media de los pesos al iniciarlo.

Al rechazar H_0 se construye el intervalo de confianza para la diferencia de medias: $IC = (\mu_1 - \mu_2 : 95\%) = [1.28; 4.72]$ esto significa que las diferencias de pesos en este tratamiento en promedio pueden variar desde 1.28 lb a 4.72 lb. El hecho de que 0 no es un valor posible en este intervalo es decir que $\mu_1 \neq \mu_2$, es otra forma de ver que podemos rechazar la hipótesis nula. Por

otro lado, analizando el valor p obtenido con el software tenemos que valor $p = 0.02$ y como $\alpha = 0.05 \geq p$ entonces llegamos a la misma conclusión, hay evidencia para rechazar H_0 .

➤ Tratamiento Familiar

Sea D = diferencia entre los pesos (en libras) de las pacientes al finalizar y al comenzar el FT. Se realizó un test similar a los anteriores para saber si con este tratamiento las pacientes aumentan de peso, entonces:

$$H_0 : \mu_1 - \mu_2 \leq 0 \quad \text{vs} \quad H_a : \mu_1 - \mu_2 > 0$$

Donde μ_1 es la media de los pesos al finalizar el Tratamiento Familiar y μ_2 es la media de los pesos al comenzar con este tratamiento. Tomando el mismo estadístico T , con los valores correspondientes a esta variable y donde s_D es el desvío de la m.a.s. $D_1; D_2; \dots; D_n$ y n es el tamaño de la muestra, y fijando $\alpha = 0.05$, realizando los cálculos manualmente, tenemos que la $RR = \{\bar{D} > 3.032\}$ y como $\bar{d} = 7.26 \in RR$ hay evidencia para rechazar H_0 . Esto significa que, en media, el peso al finalizar el tratamiento es superior que al iniciarlo. Como se rechaza H_0 se construye el intervalo de confianza para la diferencia de medias $IC = (\mu_1 - \mu_2 : 95\%) = [5.57; 8.95]$ lo que indica que esta diferencia oscila entre 5.57 lb y 8.95 lb. El hecho de que 0 no es un valor posible en este intervalo es decir que $\mu_1 \neq \mu_2$, es otra forma de ver que podemos rechazar la hipótesis nula. Por otro lado, analizando el valor p obtenido con el software tenemos que valor $p = 0.00$ y como $\alpha = 0.05 \geq p$ entonces llegamos a la misma conclusión, hay evidencia para rechazar H_0 .

➤ Comparación de tratamientos

Como las pacientes sin tratamiento (Control) no aumentaron de peso, solamente se compararán las que hicieron el Tratamiento Cognitivo con las que hicieron el Tratamiento Familiar para ver con cuál tratamiento aumentaron más de peso.

Dado que las varianzas son desconocidas debemos realizar un test de hipótesis para decidir si son iguales o no. Los datos muestrales registran una mayor variabilidad de CBT con TF.

$$CBT = \sigma_1^2 \quad m = 29$$

$$FT = \sigma_2^2 \quad n = 17$$

Vamos a poner a prueba: $H_0 : \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \quad \text{vs} \quad H_a : \frac{\sigma_1^2}{\sigma_2^2} > 1$

El pivote adecuado es: $T = \frac{s_1^2}{s_2^2} \sim F(28, 16)$ Fijando $\alpha = 0,05$, con la ayuda del software calculamos

$F(28, 16)$. Entonces, bajo H_0 , tenemos que $RR = \left\{ \frac{s_1^2}{s_2^2} > 2,02 \right\}$. En base a los datos muestrales

tenemos: $\frac{s_1^2}{s_2^2} = \frac{7,29^2}{7,16^2} = 1,03 \notin RR$, donde $s_1 = 7,29$ y $s_2 = 7,16$. Por lo tanto, no hay suficiente evidencia para rechazar H_0 , por lo que podemos suponer que las varianzas son iguales.

Llamando X = diferencia de peso (en lb) al finalizar el tratamiento CBT respecto del peso al iniciarlo, Y = diferencia de peso (en lb) al finalizar el tratamiento FT respecto del peso al iniciarlo,

$$H_0 : \mu_X - \mu_Y \geq 0 \quad \text{vs} \quad H_a : \mu_X - \mu_Y < 0$$

Usamos como pivote a $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{29} + \frac{1}{17}}} \sim t(29 + 17 - 2)$ donde $s_p^2 = \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}$ se tiene que, bajo

H_0 , $RR = \{\bar{X} - \bar{Y} < -3.720\}$. Calculando con los datos muestrales: $\bar{X} - \bar{Y} = -4,25$. Luego,

$-4,25 \in RR$, y por lo tanto, existe suficiente evidencia para rechazar H_0 , lo que significa que con el CBT aumentan menos de peso que con el FT. Como se rechaza H_0 se construye el intervalo de confianza para la diferencia de medias para muestras aleatorias independientes con varianzas desconocidas iguales $IC = (\mu_1 - \mu_2 : 95\%) = [-8.72; 0.22]$. Esto significa que la media de las diferencia de pesos oscila entre -8,72 lb y 0.22 lb. Por otro lado, analizando el valor p obtenido con el software tenemos que valor $p = 0.03$ y como $\alpha = 0.05 \geq p$ entonces llegamos a la misma conclusión, hay evidencia para rechazar H_0 .

Discusión

Al buscar información sobre los distintos tratamientos para la AN, encontramos que no había muchas certezas sobre la efectividad de éstos. Aún así, también hallamos que uno de los métodos más aceptados es el CBT, lo que también fue confirmado por la Dra. Matzkin durante la entrevista pero bajo la aclaración de que también suele recomendarse como parte de una serie de tratamientos, que también podrían incluir la FT. Una de nuestras mayores dificultades fue encontrar qué era un tratamiento efectivo, o alguna forma de poder determinar cuándo un tratamiento era efectivo o no. Al no obtener respuestas satisfactorias por parte de la bibliografía consultada o la entrevista con la Dra. Matzkin, decidimos tomar como efectivo el que aumenten de peso. Una vez definido esto, y con la información recolectada formulamos dos hipótesis.

Durante el análisis descriptivo, al observar los histogramas notamos que, las pacientes sin tratamiento bajaban de peso y que con quienes recibían FT aumentaron más de peso, como aún no podemos confirmar que realmente sea así, decidimos poner a prueba nuestras hipótesis. Pensamos primero en verificar con qué tratamientos aumentaban realmente de peso. Al realizar los respectivos test de hipótesis, encontramos que con el CONT, en promedio no lo hacían, lo que comprueba nuestra segunda hipótesis.

Luego analizamos que sucedía con CBT y FT, y como era de esperarse, ya que en los diferentes textos consultados mencionan mejoras en los pacientes bajo estos tratamientos, resultó que en promedio sí aumentaban de peso. Aún así todavía no podemos concluir si aumentan más con uno o con otro.

Por este motivo, realizamos un nuevo test de hipótesis, pero esta vez, comparando los promedios de las diferencias de peso según CBT y FT. Los resultados rechazan nuestra hipótesis, lo que nos dice que las pacientes en promedio aumentaron más de peso con la FT que con CBT.

Esto podría compararse con lo leído sobre este tratamiento, que suele recomendarse a pacientes jóvenes, más precisamente adolescentes, permitiendo la construcción de una alianza con la familia, y donde cada miembro tiene el papel de coterapeuta, fundamental en el avance de dicho tratamiento (Morandé M., Graell, S.F., n.d.).

Conclusiones

El objetivo de este estudio consistió en probar si las pacientes que padecían Anorexia Nerviosa subían de peso con el Tratamiento Cognitivo, con el Tratamiento Familiar y sin tratamiento (control) analizando también con cuál de ellos se sube más de peso. Luego de distintos análisis realizados se puede concluir, con un nivel de significación del 5%, que las pacientes aumentan más de peso con el Tratamiento Familiar que con el Tratamiento Cognitivo y que sin tratamiento no aumentan de peso.

Podemos decir que nos resultó interesante el tema para trabajar en el secundario y realizar una investigación con datos de nuestra cercanía geográfica y considerando otros factores como el índice de masa corporal, la altura, la edad de inicio del trastorno, características genéticas de sus padres,

entre otras.

Referencias

- Borrego Hernando O (2000). *Tratamiento de los trastornos de la conducta alimentaria: Anorexia y Bulimia*. Disponible en <https://goo.gl/NKuvXo> [10 de Noviembre de 2016]
- Devore, J. L. (2008). *Probabilidad y Estadística para ingeniería y ciencias*. Thomson, México DF, México, séptima edición. Disponible en <https://goo.gl/d8vDkC> [26 de Noviembre de 2016]
- Hand, DJ., Daly, K., Lunn, F., McConway, D. and Ostrowski, E. (1993) A Handbook of Small Data Sets. Chapman & Hall, Data set 285 (p. 229). Disponible en <http://www.stat.ncsu.edu/research/sas/sicl/data/> [9 de Noviembre de 2016]
- Ley N° 26.396. (2008). *Ley de trastornos alimentarios*. Disponible en <https://goo.gl/87RciT> [9 de Noviembre de 2016]
- Morandé, G., Graell, M. (n.d.). *Tratamiento psicoterapéutico de la anorexia y la bulimia nervosa*, Cap 18. Disponible en <https://goo.gl/JRmPSU> [12 de Noviembre de 2016]
- 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/>
- Saldaña, C. (2001). *Tratamientos psicológicos eficaces para trastornos del comportamiento alimentario*. Psicothema.com Disponible en <https://goo.gl/Lri2zp> [10 de Noviembre de 2016]

LA INFLUENCIA DEL TABACO EN EL PESO DE LOS RECIÉN NACIDOS

Gastón Alberto Galván Juárez - Melani Denise Mazza Caligaris

gastongalvan084@gmail.com - melani.caligaris@gmail.com

Resumen

En este trabajo se encuentra una investigación cuyo objetivo es analizar la relación entre el peso del recién nacido con el consumo de tabaco de la madre durante el embarazo. Para ello, se recogieron datos de 189 madres que se atendieron en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts, durante 1986 y se realizó la comparación de la variabilidad de los pesos de los recién nacidos teniendo en cuenta si la madre fumó durante el embarazo. Este análisis nos permitió observar si el hábito de fumar de la madre afecta al peso del bebé.

Palabras Claves bajo peso, tabaco, embarazo.

Introducción

El bajo peso¹ de los recién nacidos (BPN) es una de las mayores causas de muerte de niños menores de 1 año. Según el diario La Nación (2004), en la Argentina un poco menos del 8 % de los nacidos tiene bajo peso. Esta cifra representa más del 50% de las muertes de los bebés menores de 1 año.

Se dice que un bebé nace con bajo peso si al momento del parto pesa menos de 2500 gramos. (Infogen, 2016).

“El peso del bebé al nacer está determinado por muchos factores. Algunos de ellos son hereditarios y otros son por gran variedad de condiciones no genéticas entre los que pueden mencionarse los hábitos y el estado de salud de la madre gestante” (Hosmer & Lemeshow, 1989).

El nacimiento de niños con bajo peso podremos evitarlo con la identificación y el correcto manejo de los factores de riesgo.

“Los bebés de las madres expuestas a los cigarrillos tienen más probabilidades de nacer con bajo peso. Se ha demostrado que los hijos de las mujeres fumadoras pesan al nacer un promedio de 200 a 300 gramos menos que los descendientes de mujeres no fumadoras” (Infogen, 2016).

El objetivo de este trabajo fue analizar la relación entre el peso del recién nacido con el consumo de tabaco de la madre, durante el embarazo.

Metodología

El estudio se realizó con una muestra aleatoria simple de 189 madres, con datos recolectados en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts durante 1986, tales como si la madre fumó durante el embarazo y el peso del bebé al nacer, en gramos (Hosmer & Lemeshow, 1989).

Se efectuó un análisis descriptivo sobre el peso de los bebés recién nacidos, los cuales fueron clasificados teniendo en cuenta si la madre del niño fumó o no durante su embarazo, obteniéndose

¹ Aunque peso y masa son dos conceptos y magnitudes físicas muy diferentes, en este trabajo utilizaremos estas palabras como sinónimo ya que así se utiliza en el uso cotidiano.

En física clásica, el peso es una medida de la fuerza gravitatoria que actúa sobre un objeto. La masa, en cambio, es una propiedad intrínseca del mismo, la cantidad de materia, independiente de la intensidad del campo gravitatorio y de cualquier otro efecto.

las dos variables con las que finalmente trabajamos: “Peso del recién nacido (en gramos) con madre fumadora atendido en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts, durante 1986” y “Peso del recién nacido (en gramos) con madre no fumadora atendido en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts, durante 1986”.

Luego, se realizó el análisis inferencial.

Se utilizó el software estadístico R a través de su interfaz gráfica R Commander (versión 3.3.2).

Para llevar a cabo el objetivo, primero se utilizó una prueba chi cuadrado de bondad de ajuste con un nivel del 5% para poner a prueba la normalidad de las variables analizadas.

Además se realizó un test de hipótesis para la diferencia de medias de las poblaciones independientes, con un nivel del 5%.

Resultados y Discusiones

A continuación se muestran los resultados obtenidos a partir del análisis descriptivo realizado sobre las muestras correspondientes a cada una de las variables de estudio y luego a la comparación realizada entre ellas.

Las variables consideradas fueron:

X = Peso del recién nacido (en gramos) con madre fumadora atendido en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts, durante 1986.

Y = Peso del recién nacido (en gramos) con madre no fumadora atendido en el *Baystate Medical Center* de la ciudad de Springfield, Massachusetts, durante 1986.

Media	Desviación Estándar	CV
2771.919	659.6349	0.2379705

Tabla 1. Medidas de resumen de la variable X.

Media	Desviación Estándar	CV
3055.783	752.6769	0.2463123

Tabla 2. Medidas de resumen de la variable Y.

	Intervalo	Cantidad de datos	Porcentaje
media $\pm$ 1 desvío	(2303.1061 ; 3808.4599)	76	65 %
media $\pm$ 2 desvío	(1551.4292 ; 4561.1368)	110	94 %
media $\pm$ 3 desvío	(797.7523 ; 5313.8137)	117	100 %

Tabla 3. Prueba empírica para la variable X.

	Intervalo	Cantidad de datos	Porcentaje
media $\pm$ 1 desvío	(2112.2841 ; 3431.5539)	52	70 %



media $\pm$ 2 desvío	(1452.6492 ; 4091.1888)	71	96 %
media $\pm$ 3 desvío	(793.0143 ; 4750.8237)	73	99 %

Tabla 4. Prueba empírica para la variable Y.

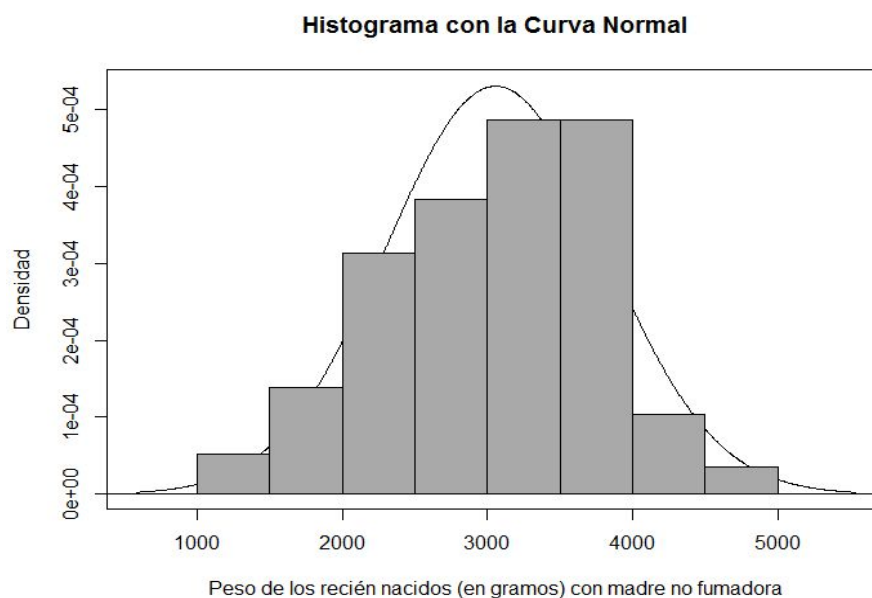


Figura 1. Histograma con Curva de Gauss de la variable “Peso de los recién nacidos con madre no fumadora (en gramos)”

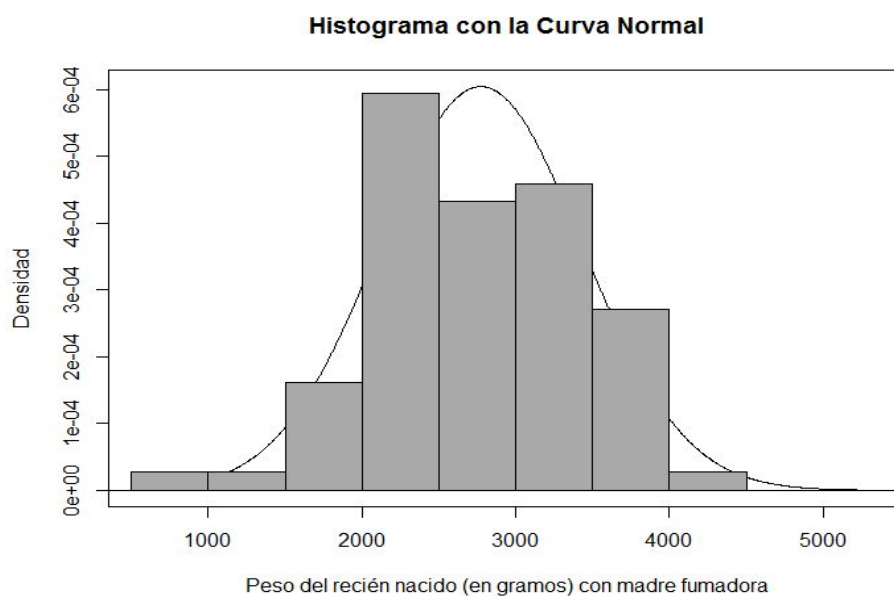


Figura 2. Histograma con Curva de Gauss de la variable “Peso de los recién nacidos con madre fumadora”

Boxplot del peso del recién nacido cuya madre es no fumadora

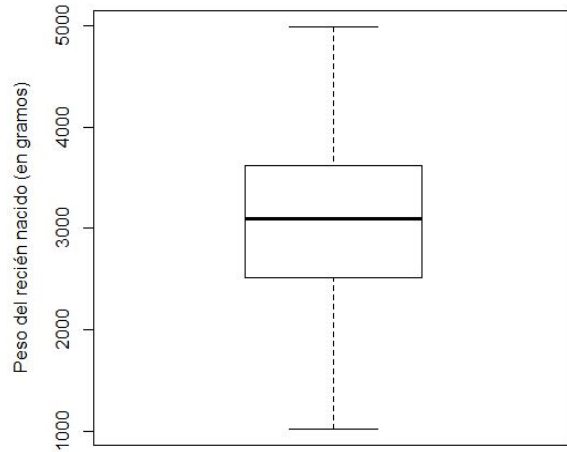


Figura 3. Boxplot de la variable “Peso del recién nacido con madre no fumadora (en gramos)”.

Boxplot del peso del recién nacido cuya madre es fumadora

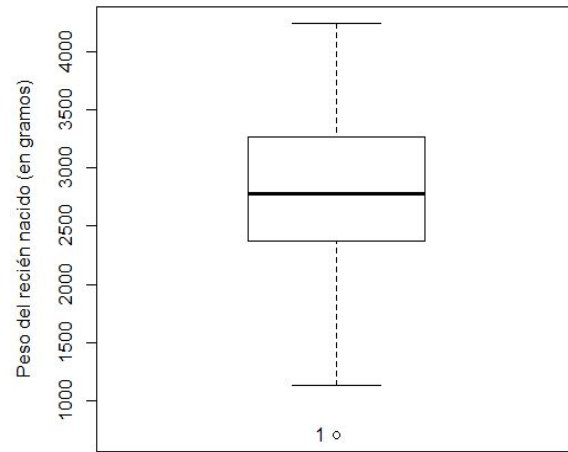


Figura 4. Boxplot de la variable “Peso del recién nacido con madre fumadora (en gramos)”.

Q-Plot del peso de los recién nacidos cuya madre es no fumadora

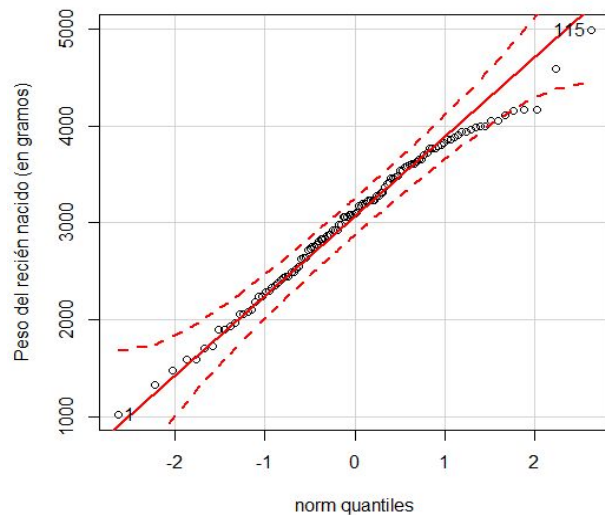


Figura 5. QQ-Plot de la variable “Peso del recién nacido con madre no fumadora”

QQ-Plot del peso de los recién nacidos cuya madre es fumadora

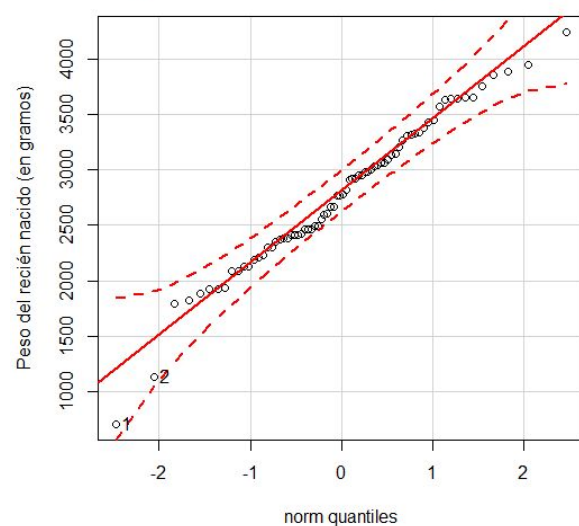


Figura 6. QQ-Plot de la variable “Peso del recién nacido con madre fumadora”

Analizando los resultados de las Tablas 1 y 2, podemos observar que las medias son las medidas de tendencia central representativas para esta muestra, ya que no se observa una importante dispersión de los datos, tal como lo afirman sus medidas de dispersión y coeficiente de variación, que se considera aceptable para un valor de aproximadamente del 30%.

En los histogramas (Figuras 1 y 2), los boxplot (Figuras 3 y 4), los QQ-Plot (Figuras 5 y 6) y las pruebas empíricas (Tablas 3 y 4), se observa que los datos aparentan estar distribuidos

normalmente.

Para analizar la distribución de las variables se utilizó un test de bondad de ajuste ($T = \sum_{i=1}^6 \frac{(o_i - e_i)^2}{e_i}$) con distribución chi cuadrado con 3 grados de libertad y el valor $p = 0.09962354$ para la variable “Peso de los bebés recién nacidos con madre no fumadora”; y $T = \sum_{i=1}^5 \frac{(o_i - e_i)^2}{e_i}$ con distribución chi cuadrado con 2 grados de libertad y el valor $p = 0.05385285$ para la variable “Peso de los bebés recién nacidos con madre fumadora”). Tomando $\alpha = 0.05$, como $\alpha < \text{valor } p$, en ambos casos no se rechazan las hipótesis nulas ($H_0: X \sim N(\mu_X, \sigma_X^2)$ y $H_0: Y \sim N(\mu_Y, \sigma_Y^2)$, respectivamente).

Dado que los parámetros poblacionales eran desconocidos, se estimaron a partir de la muestra, con $\hat{\mu}_X = \bar{X}$, $\hat{\sigma}_X^2 = S_X^2$, $\hat{\mu}_Y = \bar{Y}$ y $\hat{\sigma}_Y^2 = S_Y^2$, respectivamente. Estos estimadores son: insesgados, suficientes y eficientes (IMVU). Los estimadores utilizados para las medias ($\hat{\mu}_X$ y $\hat{\mu}_Y$) fueron hallados por el método de máxima verosimilitud. Los estimadores $\hat{\sigma}_Y^2$ y $\hat{\sigma}_X^2$ fueron hallados por el mismo método pero corregidos para que cumplan con las propiedades anteriormente nombradas (insesgados, suficientes y eficientes). Con dichas correcciones se obtuvieron $\hat{\sigma}_X^2 = S_X^2$ y $\hat{\sigma}_Y^2 = S_Y^2$.

Con respecto al análisis de la relación entre las variables, se puede observar en las Tablas 1 y 2 que los pesos de los bebés con madres no fumadoras son, en media, más elevados que los pesos de los recién nacidos de las madres que consumen tabaco (la media de la variable “Peso de los recién nacidos con madre fumadora” fue de 2771.919 gramos mientras que para la variable “Peso de los recién nacidos con madre no fumadora”, se obtuvo 3055.696 gramos de media).

Se pretende determinar si los bebés de madres no fumadoras nacen con pesos más elevados que los bebés cuyas madres han fumado durante su embarazo. Para esto se ponen a prueba las hipótesis $H_0: \mu_X - \mu_Y \geq 0$ vs $H_a: \mu_X - \mu_Y < 0$ ($G = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{s_p \cdot \sqrt{\frac{1}{n} + \frac{1}{m}}}$ con distribución t de Student, con 187 grados de

libertad y el valor $p = 0.00494546$). Para ello, se realizó un test de hipótesis para la diferencia de medias de poblaciones normales independientes, donde antes se tuvo que realizar otro test para saber si las varianzas (desconocidas) eran iguales o distintas.

Los datos muestrales ofrecen suficiente evidencia para rechazar H_0 , tomando $\alpha = 0.05$, pues $\alpha > \text{valor } p$. Por lo tanto, los valores de la variable Y son, en media, mejores que los de la variable X. Esto quiere decir que los pesos de los bebés recién nacidos disminuyen cuando la madre fuma durante el embarazo.

Según Infogen (2016), los hijos de las mujeres fumadoras pesan al nacer un promedio de 200 a 300 gramos menos que los descendientes de mujeres no fumadoras.

Esto coincide con los resultados obtenidos en este trabajo, en donde los pesos de los bebés recién nacidos en el *Baystate Medical Center* de la ciudad de Springfield son inferiores cuando la madre consume tabaco.

Conclusiones

A través de diferentes análisis se llegó a la conclusión, con un nivel de confianza del 95 %, que los pesos de los bebés recién nacidos en el *Baystate Medical Center* de la ciudad de Springfield son inferiores si su madre tuvo el hábito de fumar durante el embarazo.

Referencias

- Dieser, MP. (2014). *Estadística para futuros matemáticos. Nota de Clases*. Estadística. [material no publicado]. Facultad de Ciencias Exactas y Naturales. Universidad Nacional de La Pampa. [21 de Noviembre de 2016].
- Dieser, MP, Caverio, L., Funkner, S. (2016). *Estimadores de la media de una población con distribución normal*. Estadística. [material no publicado] Facultad de Ciencias Exactas y Naturales. Universidad Nacional de La Pampa. [21 de Noviembre de 2016].
- Hosmer, D., y Lemeshow, S. (1989). *Regresión logística aplicada*. Wiley series en probabilidad y estadística matemática. Nueva York: Wiley.
- Infogen (2016). *Defectos al nacimiento*. [online] Disponible en <http://infogen.org.mx/bajo-peso-al-nacer/> [21 de Noviembre de 2016].
- La Nación (2004). *Nacen con menos peso los bebés argentinos*. [online] Disponible en <http://www.lanacion.com.ar/644791-nacen-con-menos-peso-los-bebes-argentinos>. [21 de Noviembre de 2016].
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria (versión 3.3.2) [Software]. Disponible en <http://www.R-project.org/>
- StatCrunch (2012). *Estudio del bajo peso al nacer*. [online] Disponible en <http://www.statcrunch.com/5.0/shareddata.php?keywords=weighing> [21 de Noviembre de 2016].
- Peso (2016). En *Wikipedia. La enciclopedia libre*. Disponible en <https://es.wikipedia.org/wiki/Peso> [21 de Noviembre de 2016].

Análisis de la relación entre el nivel de colesterol y el infarto de miocardio

Dahiana Gimenez - Jorgelina Blanco - Laura Caffarone

dahiana364@gmail.com - blancojorgelina25@gmail.com - mlcaffalp@gmail.com

Resumen

En el presente trabajo se pretende analizar el riesgo de sufrir un infarto agudo de miocardio a causa del colesterol. Para llevarlo a cabo tomamos los datos obtenidos a partir de un estudio realizado en un centro médico del noreste de Estados Unidos y realizamos un análisis descriptivo e inferencial para determinar posibles relaciones que existen con respecto al nivel de colesterol en pacientes que han sufrido un infarto y aquellos que no. Analizamos también cómo se comporta este nivel en pacientes con esta cardiopatía con el transcurso del tiempo.

Introducción

El colesterol es una sustancia similar a la grasa e indispensable para la vida. Se encuentra en las membranas celulares de nuestro organismo. El cuerpo necesita colesterol para fabricar hormonas, ácidos biliares, vitamina D, y otras sustancias. Sin embargo, el aumento del colesterol en la sangre y su depósito en las arterias puede ser peligroso y producir aterosclerosis (estrechamiento o endurecimiento de las arterias por depósito de colesterol en sus paredes). (Fundación Hipercolesterolemia Familiar, n.d)

El presente trabajo pretende indagar con respecto a algunas relaciones que podrían ocurrir entre el colesterol y los infartos, dado que las enfermedades cardiovasculares son la causa principal de muerte en los países occidentales, (Gutierrez Abejon, E. 2010).

El aumento en las cifras de colesterol en sangre y su depósito en las arterias origina la enfermedad aterosclerótica cardiovascular que es la principal causa de mortalidad. La enfermedad aterosclerótica incluye la enfermedad coronaria (angina de pecho e infarto de miocardio), cerebrovascular (infarto cerebral o ictus) y la enfermedad arterial periférica (claudicación o dolor en las piernas). El colesterol elevado es uno de los principales factores de riesgo cardiovascular, junto con el tabaco, la hipertensión arterial y la diabetes mellitus. (Fundación Hipercolesterolemia Familiar, n.d).

En este trabajo, la investigación estará basada en el infarto agudo de miocardio (músculos del corazón), que es la muerte del órgano o parte de él por falta de riego sanguíneo, debido a una obstrucción de la arteria correspondiente a éste.

Con un enfoque estadístico de dicho análisis utilizaremos los datos obtenidos en un estudio que se realizó en un centro médico del noreste de Estados Unidos con respecto a los niveles de colesterol en sangre a pacientes que sufrieron un infarto y a otro grupo que no, con los siguientes objetivos:

- Determinar si el nivel medio de colesterol difiere de los pacientes que sufren ataques cardíacos, de los que no.
- Determinar si el nivel de colesterol en sangre influye en la patología de infarto.
- Analizar si, después de un ataque cardíaco y en el transcurso del tiempo, el nivel de colesterol en sangre disminuye.

Para poder tener una referencia de lo que se considera nivel de colesterol deseable o alto, nos basamos en la siguiente tabla extraída de Fundación Hipercolesterolemia Familiar. Colesterol y Triglicérido (n.d).

Valores normales y elevados del Perfil Lipídico

Colesterol total	
Por debajo de 200 mg/dl	Deseable
200-239 mg/dl	Límite alto
240 mg/dl	Alto
Por debajo de 180 mg/dl (menor de 18 años)	Deseable

Tabla 1: Clasificación según el nivel de colesterol.

Metodología

La población a analizar en esta investigación está compuesta por pacientes que han sufrido un ataque cardíaco y pacientes de un grupo de control que no han sufrido dicha cardiopatía, en un centro médico del noreste de los Estados Unidos.

La muestra que utilizamos consta de un total de 28 pacientes, que sufrieron ataque cardíaco, a los que se le midió su nivel de colesterol en sangre a los 2 y 14 días después del mismo. Además, se registraron los niveles de colesterol de un grupo control de 30 personas que no habían tenido esta cardiopatía, suponiendo que los datos observados son de individuos adultos.

Las unidades de medida de colesterol no se dan en la referencia original, pero son probablemente mg/dl de sangre (Ryan *et al.*, 1985).

Las variables utilizadas en nuestra investigación fueron:

1. El nivel de colesterol en sangre de un paciente de un centro médico del noreste de los EEUU que sufrió un infarto. (En este caso tomamos los datos de la extracción de sangre del día 2, pues es la más próxima al infarto y la extracción del día 14, ya que es la más distanciada de dicha cardiopatía).
2. El nivel de colesterol en sangre de pacientes de un centro médico del noreste de EEUU que no sufrieron infarto.
3. Además, utilizamos una variable derivada de la primera, la cual definimos como la diferencia entre el nivel de colesterol a los dos días del infarto y a los catorce días del mismo.

Para trabajar con todos los datos ofrecidos en esta muestra, se realizó una depuración de éstos, pues se observó que para el día 14, en nueve casos, no se habían registrado los niveles de colesterol de las extracciones correspondientes. Para completar dichos datos se calculó la media muestral (221.47) y se completaron cinco observaciones con 221 y los cuatro faltantes con 222 para obtener un valor aproximado de la media. Por otra parte, teníamos otro grupo de datos (filas del día 4 y columna grupo) innecesarios para nuestro análisis, los cuales fueron suprimidos. Se utilizó el software R (versión 3.3.0) para realizar la depuración de datos, así como también para construir gráficos y hacer cálculos auxiliares.

Las variables aleatorias que se han trabajado están relacionadas con los niveles de colesterol en

sangre. La mayoría de las dimensiones del cuerpo humano, como la mayoría de los fenómenos naturales, se distribuyen normalmente. Muchas variables aleatorias continuas presentan una función de densidad cuya gráfica tiene forma de campana.

Es de esperar que, en una población razonablemente homogénea, la distribución de cualquiera de sus dimensiones antropométricas sea normal, y por ello, las estimaciones, cálculos y, en general cualquier tratamiento estadístico, puede efectuarse según las propiedades de esta distribución, lo que es muy conveniente dada la facilidad que el tratamiento de esta distribución supone (Valero Cabello, n.d). Para poder verificar esta hipótesis de manera estadística realizamos una prueba de bondad de ajuste para cada variable utilizada (ver Anexo A).

Con respecto a la comparación de los niveles medios de colesterol en sangre de los pacientes que sufrieron un infarto con los que no, reconocimos que ambas variables utilizadas eran independientes pues se está trabajando con dos grupos de individuos diferentes. Lo primero que se utilizó fue un

Test de hipótesis para analizar la igualdad de varianzas, usando el estadístico $T = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F(n-1; m-1)$

(Dieser, 2016). Y después se realizó la comparación de las medias con un Test de hipótesis utilizando el estadístico $T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n} + \frac{s_2^2}{m}\right)}} \sim t(\eta)$ (Dieser, 2016). Luego para la diferencia de niveles

medios de colesterol en los días 2 y 14 del mismo paciente se realizó un Test de hipótesis usando el estadístico $T = \frac{\bar{D} - (\mu_1 - \mu_2)}{\frac{s_D}{\sqrt{n}}} \sim t(n - 1)$ (Dieser, 2016). En todos los casos definimos un nivel de

significación de la prueba de 0,01, pues consideramos el menor valor para que el error sea lo más serio posible.

Finalmente para determinar si el nivel de colesterol en sangre influye en aquellos que sufren infarto, analizamos un gráfico de boxplot y la Tabla 1 de niveles de colesterol.

Resultados

En primer lugar construimos un boxplot (Figura 1 y 2) para cada una de las siguientes variables aleatorias independientes, nivel de colesterol en sangre (en mg/dl) de pacientes a los 2 días de sufrir la cardiopatía ($\bar{x} = 253.93$ y $s_1 = 47.71$) y el nivel de colesterol en sangre (en mg/dl) de pacientes de control ($\bar{y} = 193.13$ y $s_2 = 22.3$). Para poder realizar un análisis más profundo de estas variables realizamos un test de hipótesis. Ya que las varianzas poblacionales son desconocidas, tuvimos que analizar si estas últimas eran iguales o diferentes, concluyendo que eran distintas (ver anexo B). Finalmente pudimos comprobar que el nivel de colesterol medio de los pacientes a los 2 días del infarto, no solo es diferente al nivel de colesterol medio de aquellos que no sufrieron ataques cardíacos, sino que además es mayor, donde $\bar{x} - \bar{y} = 60.8$ que pertenece a $RR = \{\bar{X} - \bar{Y} > 24.04\}$ (ver anexo B).

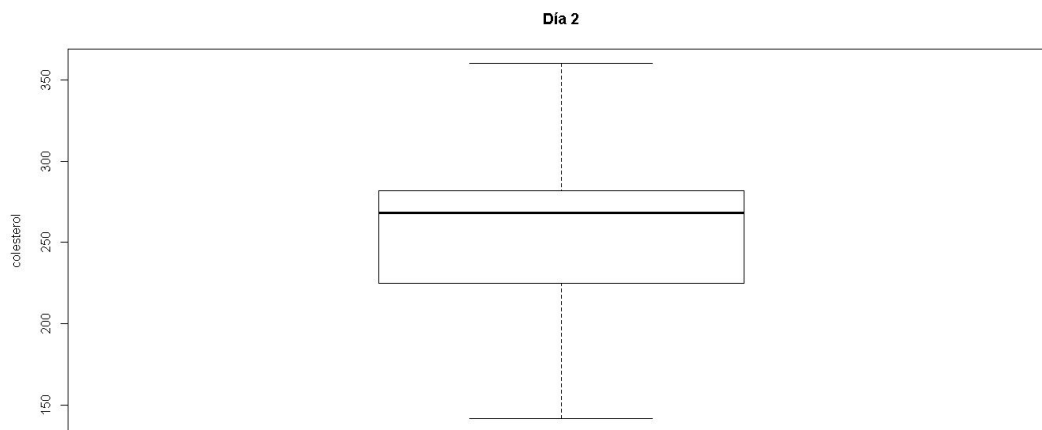


Figura 1: Boxplot del nivel de colesterol en sangre (mg/dl) de 28 pacientes a los 2 días de sufrir un infarto.

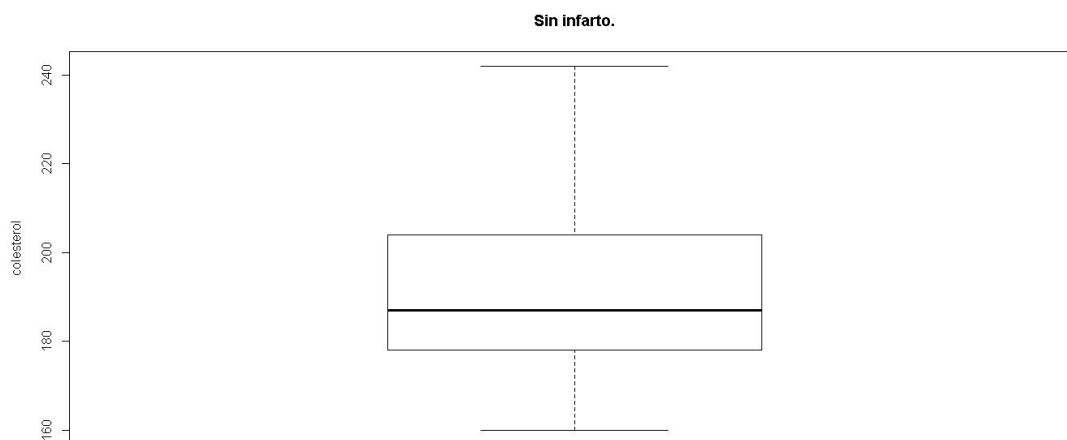


Figura 2: Boxplot del nivel de colesterol en sangre (mg/dl) de 30 pacientes que no sufrieron infarto.

Se construyeron dos histogramas de frecuencias observadas y se realizó un test de hipótesis para variables aleatorias dependientes, a partir de las cuales elaboramos una variable aleatoria que la denominamos diferencia de los niveles de colesterol en sangre (en mg/dl) en la extracción del día 2 y del día 14 después del infarto de un paciente de un centro médico del noroeste de los EEUU. Y utilizando un nivel de significación de 0,01 pudimos determinar que, como la muestra arroja una media muestral de 43.14 , que pertenece a $RR = \{D > 18,15\}$ elaborada, entonces podemos afirmar que hay suficiente evidencia para rechazar la hipótesis nula, es decir, que el nivel medio de colesterol de las extracciones del día 2 es mayor a las del día 14. (Para observar todo el procedimiento ver anexo C)

A continuación se muestran los gráficos obtenidos.

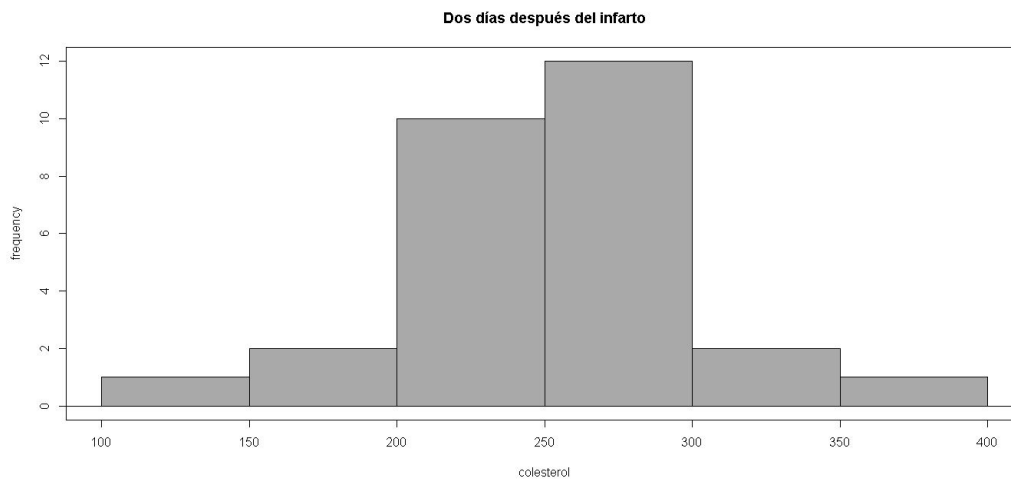


Figura 3: Histograma del nivel de colesterol en sangre en 28 pacientes, dos días después del infarto.

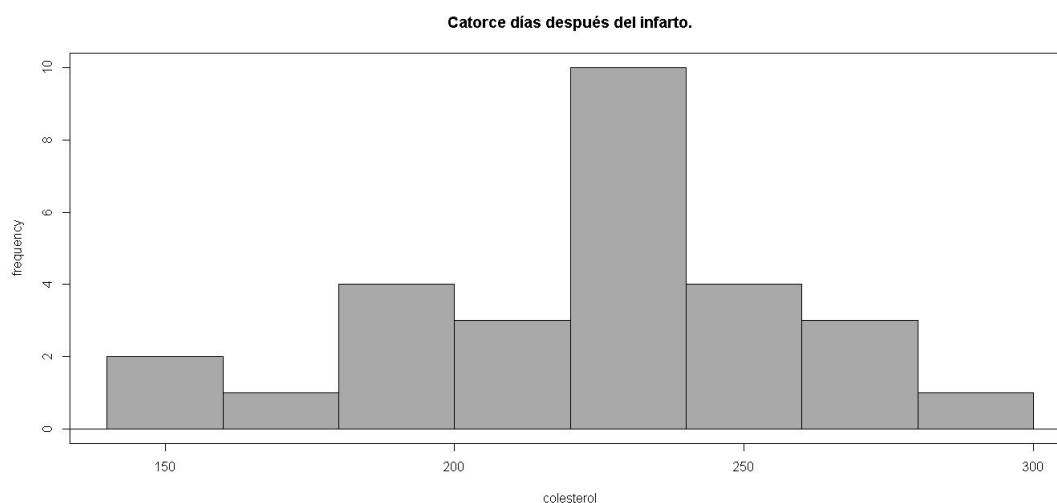


Figura 4: Histograma del nivel de colesterol en sangre en 28 pacientes, catorce días después del infarto.

Discusión

Comparando los histogramas (Figuras 3 y 4), podemos ver que hay más frecuencias observadas en niveles de colesterol más altos en el día 2 después del infarto, donde la mayor cantidad de observaciones se encuentran entre los 200 mg/dl y 300 mg/dl, teniendo datos hasta los 400 mg/dl de colesterol en sangre. En cambio, en las frecuencias observadas del día 14, se puede ver que éstas están más dispersas, tomando valores desde los 150 mg/dl a 300 mg/dl de colesterol en sangre y donde el máximo de las frecuencias observadas se produce en niveles de colesterol menores a 250 mg/dl en sangre. Por lo tanto, podemos decir que el análisis inferencial que realizamos en nuestra investigación confirma el análisis que obtuvimos con los gráficos.

Por otra parte, los diagramas de caja y de brazos permiten comparar la distribución del nivel de colesterol en sangre. De esta manera observamos que el nivel medio de colesterol de los pacientes que han sufrido infarto es más alto, mientras que el nivel medio de colesterol de los que no sufrieron esta cardiopatía están más concentrados entre 150 mg/dl y 200 mg/dl. Así mismo se puede observar que en el primer caso (Figura 1), el máximo se encuentra por encima de los 350 mg/dl, y

en el segundo caso (Figura 2) se encuentra en 240 mg/dl aproximadamente.

Tomando la información precedente y observando la información de la Tabla 1, podemos decir, que el nivel medio de colesterol en sangre de los pacientes que han sufrido un infarto, no solo es superior a los que no lo han sufrido, sino que además según la referencia de la tabla corresponde a un nivel alto, mientras que en los pacientes de control, el nivel de colesterol pertenece a un nivel deseable.

Conclusiones

Los objetivos de este trabajo se centraron en determinar si el nivel medio de colesterol difiere de los pacientes que sufren un ataque de miocardio de los que no, determinar si el nivel de colesterol influye en el mismo y analizar si después de un ataque cardíaco y en el transcurso del tiempo, el nivel de colesterol en sangre disminuye. Luego de los análisis realizados, descriptivo e inferencial, logramos corroborar en base a la evidencia aportada por las muestras que: el nivel medio de colesterol de los pacientes que sufrieron un ataque cardíaco difiere y es mayor al nivel medio de colesterol en pacientes que no sufrieron un ataque de miocardio. Por otro lado, determinamos que el nivel de colesterol alto influye en los ataques cardíacos, y además observamos que con el paso de los días, el nivel de colesterol en los pacientes con ataque cardíaco fue disminuyendo.

Referencias

- Dieser, MP. (2016). *Estadística para futuros matemáticos: Notas de clase*. Material no publicado. FCEyN-UNLPam. Santa Rosa-La Pampa.[5 de noviembre de 2016]
- Fundación Hipercolesterolemia Familiar (n.d). *Colesterol y Triglicéridos*. [online] Disponible en: (<https://goo.gl/FCSfiQ>) [7 de noviembre de 2016]
- Gutierrez Abejon, E. (2010). La aterosclerosis como desencadenante de la patología cardiovascular. *Biociencias*, 7, 1-28. Disponible en: <http://www.uax.es/publicacion/la-aterosclerosis-como-desencadenante-de-la-patologia-cardiovascular.pdf>
- Ryan, BF., Joiner, BL. & Ryan Jr, TA. (1985). *Minitab Handbook Second Edition*. Boston: Duxbury Press.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [software] Disponible en: <https://www.R-project.org/>.
- Valero Cabello, E.(nd). *Antropometría*. Centro Nacional de Nuevas Tecnologías Instituto Nacional de Seguridad e Higiene en el Trabajo. [online] Disponible en: <https://goo.gl/2hR9lp> [7 de noviembre de 2016]

Anexo A:

Sea la variable aleatoria X = nivel de colesterol en sangre en pacientes a los 2 días de sufrir un infarto.

Tomando 9 clases y dado que los parámetros poblacionales son desconocidos, podemos estimarlos a partir de la muestra, de manera que $\hat{\mu} = \bar{x} = 253.93$ y $\hat{\sigma}^2 = s^2 = (47.71)^2$. Por lo tanto las hipótesis a poner a prueba son:

$$H_0: X \sim N(253.93; (47.71)^2) \quad \text{Vs.} \quad H_a: X \text{ no sigue la distribución propuesta.}$$

Bajo la hipótesis nula, los valores observados y esperados se muestran en el Cuadro 1.

i	clase	o_i	e_i
1	$(-\infty; 180]$	2	1.68
2	$(180; 200]$	1	1.96
3	$(200; 220]$	3	4.76
4	$(220; 240]$	5	6.16
5	$(240; 260]$	2	9.24
6	$(260; 280]$	7	10.36
7	$(280; 300]$	5	12.88
8	$(300; 320]$	2	12.88
9	$(320; \infty)$	1	2.35

Cuadro 1: Frecuencias observadas y esperadas de colesterol en sangre a los 2 días de sufrido un infarto.

Los valores del Cuadro 1 han sido calculados con el software R (versión 3.3.0), considerando la probabilidad de ocurrencia de cada categoría y el número total de observaciones, en este caso, 28 pacientes. Notamos que existen frecuencias esperadas inferiores a 5, por lo tanto es necesario agrupar las clases. Luego se obtiene la agrupación que se muestra en el cuadro 2.

i	clases	o_i	e_i	$\frac{(o_i - e_i)^2}{e_i}$
1	$(-\infty; 220]$	6	8.4	0.68
2	$(220; 260]$	7	8.48	0.26
3	$(260; 300]$	12	7.64	2.49
4	$(300; \infty)$	3	5.15	0.9

Cuadro 2: Frecuencias observadas, esperadas y desvíos de nivel de colesterol en sangre a los 2 días de sufrido un infarto.

El estadístico de prueba resulta: $T = \sum_{i=1}^6 \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(1)$

Fijando $\alpha = 0.01$ tenemos: $RR = \{T \geq \chi^2(1)_{0.01}\} = \{T \geq 6.63\}$

Por lo tanto, sumando los desvíos cuadráticos relativos de la quinta columna del Cuadro 2 se tiene $t = 4.33 \notin RR$, luego no hay suficiente evidencia al nivel del 1% para suponer que los datos no provienen de una distribución normal de parámetros 253.93 y $(47.71)^2$.

Análogamente a este procedimiento resulta el análisis de la variable aleatoria W = nivel de colesterol en sangre (en mg/dl) de pacientes a los 14 días del infarto.

De la misma manera siendo la variable aleatoria Y = nivel de colesterol en sangre en pacientes que no sufrieron infartos.

Tomando 9 clases, y dado que los parámetros poblacionales son desconocidos, podemos estimarlos a partir de la muestra, de manera que $\hat{\mu} = \bar{y} = 193.13$ y $\hat{\sigma}^2 = s^2 = (22.3)^2$. Por lo tanto las hipótesis a poner a prueba son:

$H_0: Y \sim N(193.13; (22.3)^2)$ vs. $H_a: Y$ no sigue la distribución propuesta.

Bajo la hipótesis nula, los valores observados y esperados se muestran en el Cuadro 3.

i	clase	o_i	e_i
1	$(-\infty; 170]$	5	4.5
2	$(170; 180]$	3	3.9
3	$(180; 190]$	8	4.8
4	$(190; 200]$	6	5.16
5	$(200; 210]$	2	5.1
6	$(210; 220]$	2	3
7	$(220; 230]$	1	2.4
8	$(230; 240]$	2	0.6
9	$(240; \infty)$	1	0.54

Cuadro 3: Frecuencias observadas y esperadas de colesterol en sangre de pacientes que no sufrieron infarto.

Los valores del Cuadro 4 han sido calculados con el software R (versión 3.3.0), considerando la probabilidad de ocurrencia de cada categoría y el número total de observaciones, en este caso, 30 pacientes. Notemos que existen frecuencias esperadas menores a 5, por lo tanto agrupamos las clases de la siguiente manera (Cuadro 4).

i	clases	o_i	e_i	$\frac{(o_i - e_i)^2}{e_i}$
1	$(-\infty; 190]$	16	13.2	0.59
2	$(190; 200]$	6	5.16	0.14
3	$(200; 210]$	2	5.1	1.88
4	$(210; \infty)$	6	6.54	0.044

Cuadro 4: Frecuencias observadas, esperadas y desvíos de nivel de colesterol en sangre de pacientes que no sufrieron infarto.

El estadístico de prueba resulta: $T = \sum_{i=1}^6 \frac{(o_i - e_i)^2}{e_i} \sim \chi^2(1)$

Fijando $\alpha = 0.01$ tenemos: $RR = \{T \geq \chi^2(1)_{0.01}\} = \{T \geq 6.63\}$

Por lo tanto, sumando los desvíos cuadráticos relativos de la quinta columna del Cuadro 4 se tiene $t = 2.65 \notin RR$, luego no hay suficiente evidencia al nivel del 1% para suponer que los datos no provienen de una distribución normal de parámetros 193.13 y $(22.3)^2$.

Anexo B:

Sea X = nivel de colesterol en sangre (en mg/dl) de pacientes a los 2 días de sufrir un infarto
 $X_1, X_2, \dots, X_{28}$ una m.a.s. $X \sim N(\mu_1, \sigma_1^2)$
 $\bar{x} = 253,93$ $s_1 = 47,71$

Y = nivel de colesterol en sangre (en mg/dl) de pacientes que no sufrieron infarto
 $Y_1, Y_2, \dots, Y_{30}$ una m.a.s. $Y \sim N(\mu_2, \sigma_2^2)$
 $\bar{y} = 193,13$ $s_2 = 22,3$

Por lo tanto las hipótesis a poner a prueba son:

$$H_0: \mu_1 - \mu_2 \leq 0 \quad \text{Vs.} \quad H_a: \mu_1 - \mu_2 > 0$$

¿Las varianzas son iguales o no?

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \quad \text{Vs.} \quad H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1$$

$$T = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \sim F(27, 29) \quad \alpha = 0,01$$

$$0,01 = P\left(\frac{s_1^2}{s_2^2} > F(27, 29)_{0,01}\right) \text{ bajo } H_0$$

Luego la región de rechazo (RR) resulta: $RR = \left\{\frac{s_1^2}{s_2^2} > 2,45\right\}$

$\frac{s_1^2}{s_2^2} = 4,58 \in RR$, luego hay suficiente evidencia para rechazar H_0 , por lo tanto las varianzas no son iguales.

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n} + \frac{s_2^2}{m}\right)}} \sim t(\eta) \Rightarrow \eta = \frac{\left(\frac{s_1^2}{n} + \frac{s_2^2}{m}\right)^2}{\frac{\left(\frac{s_1^2}{n}\right)^2}{n-1} + \frac{\left(\frac{s_2^2}{m}\right)^2}{m-1}} = \frac{(81,3 + 16,6)^2}{244,8 + 0,02} = 39,15$$

$$RR = \{\bar{X} - \bar{Y} > t(39)_{0,01} \sqrt{\left(\frac{(47,71)^2}{28} + \frac{(22,3)^2}{30}\right)}\}$$

$$RR = \{\bar{X} - \bar{Y} > 2,43 \cdot \sqrt{97,87}\} \Rightarrow RR = \{\bar{X} - \bar{Y} > 24,04\}$$

$\bar{x} - \bar{y} = 60,8 \in RR$, luego hay suficiente evidencia para rechazar H_0 , es decir, podemos decir que las medias son diferentes y además $\mu_1 > \mu_2$.

Anexo C:

Siendo:

X = nivel de colesterol en sangre (en mg/dl) de pacientes a los 2 días de sufrir un infarto

$X_1, X_2, \dots, X_{28}$ una m.a.s. $X \sim N(\mu_1, \sigma_1^2)$

W = nivel de colesterol en sangre (en mg/dl) de pacientes a los 14 días del infarto

$W_1, W_2, \dots, W_{28}$ una m.a.s. $W \sim N(\mu_2, \sigma_2^2)$

D = Diferencia de los niveles de colesterol en sangre (en mg/dl) en la extracción del día 2 y del día 14 después del infarto de un paciente de un centro médico del noreste de los EEUU.

Por lo tanto podemos considerar la variable aleatoria $D = X - W$, que por ser combinación lineal de normales es normal. $D_1, D_2, \dots, D_{28}$ m.a.s. de $D \sim N((\mu_1 - \mu_2), \sigma_3^2)$

Por lo tanto las hipótesis a poner a prueba son:

$$H_0: \mu_1 - \mu_2 \leq 0 \quad \text{Vs.} \quad H_a: \mu_1 - \mu_2 > 0$$

Se utilizó el siguiente estadístico $T = \frac{\bar{D} - (\mu_1 - \mu_2)}{\frac{s_D}{\sqrt{n}}} \sim t(n-1)$

$D_1=114 \ D_2=15 \ D_3=-32 \ D_4=79 \ D_5=59 \ D_6=16 \ D_7=18 \ D_8=4 \ D_9=-22 \ D_{10}=21 \ D_{11}=18 \ D_{12}=30$
 $D_{13}=-15 \ D_{14}=118 \ D_{15}=30 \ D_{16}=60 \ D_{17}=-30 \ D_{18}=2 \ D_{19}=88 \ D_{20}=100 \ D_{21}=66$
 $D_{22}=96 \ D_{23}=58 \ D_{24}=80 \ D_{25}=66 \ D_{26}=32 \ D_{27}=-36 \ D_{28}=32$

Luego $s_D = 38.85 \ \bar{d} = 43.14$

Fijando $\alpha = 0.01 \quad RR = \{ \bar{D} > t(n-1)\alpha \frac{s_D}{\sqrt{n}} \} = \{ \bar{D} > 18.15 \}$

La muestra arroja $\bar{d} = 43.14 \in RR$. Luego hay suficiente evidencia para rechazar H_0 , es decir, los niveles medios de colesterol de las extracciones de los días 2 y 14, disminuye con el paso de los días después del infarto.

RELACIÓN ENTRE DEPRESIÓN Y NIVELES DE ACTH Y CORTISOL

Angel Gabriel Martin - Eduardo Daniel Agüero

martinangelgabriel@gmail.com - eduardo.daniel.aguero@hotmail.com

Resumen

La depresión es una patología que se está viendo cada vez con más frecuencia en jóvenes y adolescentes. Es por ello que este trabajo presenta una investigación sobre esta enfermedad y su vinculación con niveles de cortisol y de hormona adrenocorticotrópica (ACTH) que produce. Observamos los niveles de cortisol y ACTH de 12 pacientes con trastornos depresivos unipolares o bipolares y 13 individuos sanos. Los resultados obtenidos a partir de las muestras evidencian una diferencia en los niveles tanto de cortisol como de ACTH. Luego de analizar la distribución de las variables y ver que se ajustan a una normal, realizamos un test inferencial para ver si estas diferencias eran significativas. Determinamos que no contamos con evidencia suficiente para asegurar que los pacientes depresivos tienen mayor nivel de cortisol y de ACTH que los sanos.

Introducción

La hipófisis es una glándula ubicada en la base del cerebro. Consta de dos partes: la adenohipófisis, formada por los lóbulos anterior e intermedio, y la neurohipófisis, que comprende el lóbulo posterior y está constituida por los axones del hipotálamo. Es en el lóbulo anterior de la adenohipófisis donde se producirá la corticotropina (ACTH) una de las encargadas de mantener metabólicamente activas a las glándulas endocrinas y estimular sus secreciones. Además, es la encargada de estimular la secreción de cortisol, una de las hormonas producidas por la corteza suprarrenal, que se libera como respuesta al estrés y a un nivel bajo de glucocorticoides en la sangre (Curtis et al., 2008). Jubiz (1993) menciona que *el síndrome de Cushing es una constelación de signos y síntomas que se deben a un exceso de cortisol*.

En una investigación, el cortisol plasmático se analizó cada 20 minutos durante 24 horas en seis pacientes psicóticamente deprimidos y ocho personas normales. Los depresivos, mientras estaban enfermos, secretaban sustancialmente más cortisol, tenían más episodios de secreción y más minutos de secreción activa (Sachar et al., 1973). A partir de la lectura de este artículo nos pareció interesante y útil enfocar esta investigación en la relación que existe entre la hormona ACTH y el cortisol en pacientes con depresión, en relación a otros pacientes sanos. Notamos además, que las investigaciones datan desde 1973 y con una gran variedad de información, lo que nos llevó a plantear los siguientes objetivos:

- Determinar si existen diferencias entre los niveles medios de cortisol en sangre en pacientes enfermos en relación a otros sanos.
- Determinar si existen diferencias en los niveles medios de ACTH en sangre en pacientes enfermos en relación a otros sanos.

Si determinamos que existe diferencia entre los niveles de cortisol en sangre en pacientes que padecen depresión unipolar o bipolar, en relación a otros sanos, así como también en los niveles de ACTH en sangre, podremos trabajar sobre estos pacientes, para prevenir en ellos estos niveles altos de cortisol o tratarlos. Con esto evitaremos que el paciente padezca las consecuencias provocadas por el síndrome de *Cushing*.

Metodología

Para este trabajo se consideró una muestra conformada por dos grupos, un primer grupo de 12 sujetos (Edad media $48,7 \pm 15,1$ años; Peso corporal medio $72,5 \pm 13,4$ kg), ingresados en la Clínica Psiquiátrica de la Universidad de Würzburg con grave depresión unipolar mayor, trastorno o depresión bipolar (según DSM-III-R²) y un segundo grupo de 13 sujetos no hospitalizados (Edad media: $47,4 \pm 6,4$ años y Peso corporal medio de $69,2 \pm 8,6$ kg).

Todos los sujetos fueron examinados cuidadosamente por estado y prevalencia de enfermedades médicas, así como abuso de alcohol o nicotina y vida estresante. Eventos que podrían influir en la hipófisis-adrenal o función del sistema inmunológico.

Se les realizaron exámenes físicos y la investigación de rutina de laboratorio, todos los pacientes estuvieron libres de drogas durante al menos 4 días antes de la investigación.

Se hicieron diagnósticos independientemente por dos psiquiatras, que eran ajenos a las investigaciones del laboratorio. La severidad de la depresión fue evaluada el día de la investigación de laboratorio utilizando la escala de 21 puntos Hamilton Rating Depresión (HRS-D, Hamilton, 1960). Los pacientes sanos no tenían historia de trastorno psiquiátrico, ni tenían parientes con una enfermedad psiquiátrica, las puntuaciones fueron inferiores a 5.

El consentimiento informado por escrito se obtuvo de todos los sujetos antes de la investigación. Este estudio fue aprobado por el comité de ética de la Universidad de Würzburg (Wodarz et al., 1991).

Primeramente se desarrolló un análisis descriptivo a través de histogramas y tablas con medidas de posición y dispersión y finalmente un análisis inferencial con un test de hipótesis para diferencia de medias de poblaciones normales. Los cuantiles de las distribuciones de los estadísticos utilizados en los test de hipótesis, así como los histogramas y las medidas de posición y dispersión se obtuvieron utilizando el software estadístico R a través de su interfaz gráfica R Commander (R Core Team, 2016). Las tablas se confeccionaron en Microsoft Office Excel (Microsoft Office Excel, 2007).

Resultados

Niveles de ACTH en pacientes con depresión y en pacientes sanos.

Analizamos, en primera instancia, si hay relación entre los niveles de ACTH en pacientes sanos y enfermos. Para hacerlo, utilizamos las muestras, y a partir de ella determinamos las medidas de resumen que se muestran en la Tabla 1.

ACTH	Media	Desvío	Mediana	MAD	n
ESTADO DEL PACIENTE					
Enfermo	4,29	1,50	4,29	1,63	12
Sano	3,57	0,94	3,96	0,98	13

Tabla 1: Medidas de resumen de la variable ACTH separadas por pacientes enfermos y sanos.

Podemos decir, a partir de la Tabla 1, que los pacientes enfermos bajo estudio tienen en promedio

² El Manual diagnóstico y estadístico de los trastornos mentales (DSM), es una obra que contiene una clasificación de los trastornos mentales y proporciona descripciones claras de las categorías diagnósticas, con el fin de que los clínicos e investigadores de las ciencias de la salud puedan diagnosticar, estudiar e intercambiar información y tratar los distintos trastornos mentales (Wikipedia, 2016).

un nivel de ACTH mayor al de los sanos; además de que su dispersión es mucho mayor. Si a su vez comparamos la $(MAD/0,6745)=2,42$ con el Desvío=1,50 en enfermos y la $(MAD/0,6745)=1,45$ con el Desvío=0,94 en sanos, podemos ver claramente que en ambos casos, en promedio, los niveles de ACTH se concentran más alrededor de la media que de la mediana. Luego continuamos nuestro análisis a partir de los histogramas de la Figura 1.

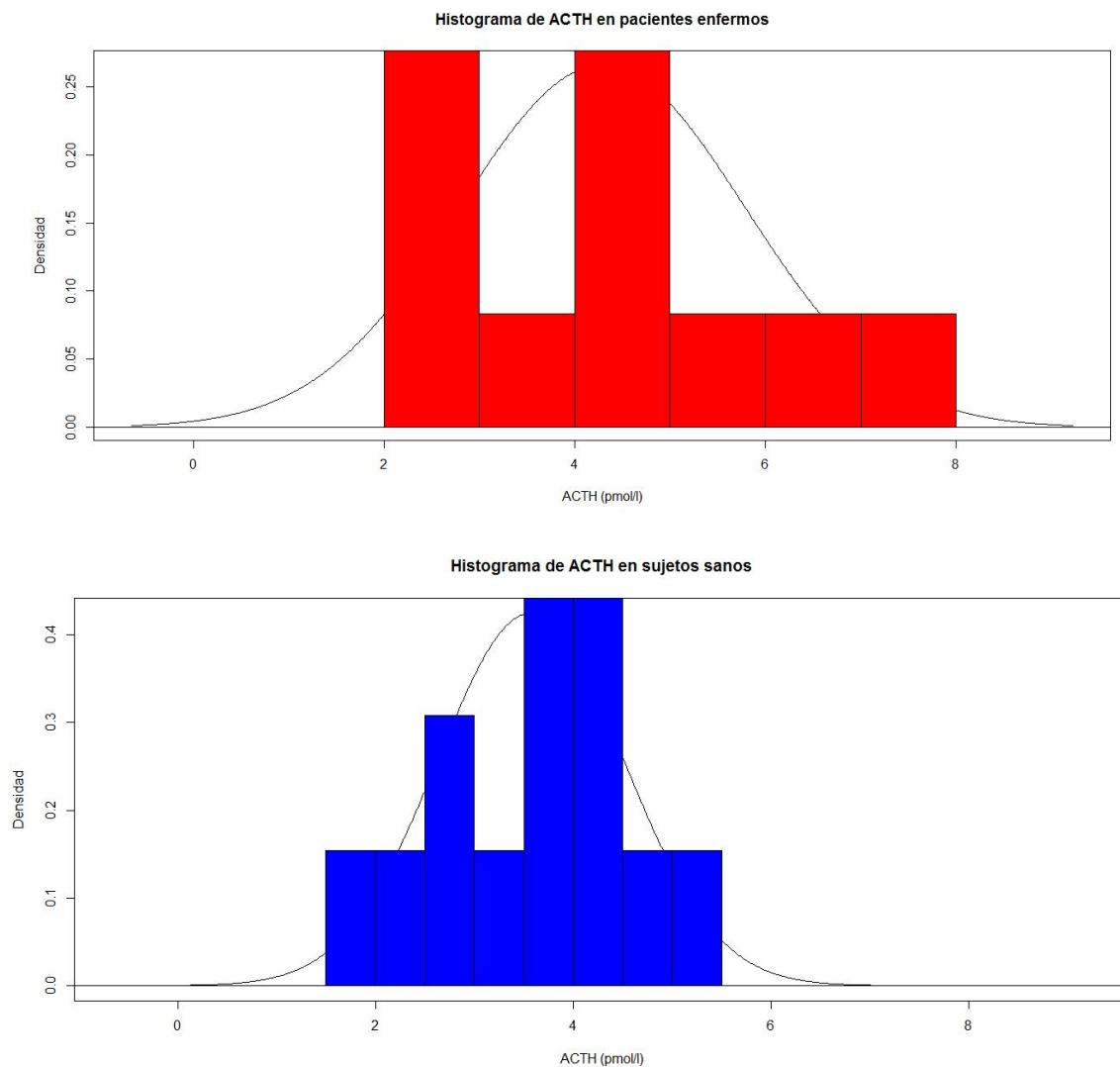


Figura 1: Histogramas de los niveles de ACTH (medidos en pmol/l) en 13 sujetos sanos y 12 enfermos (con depresión unipolar o bipolar) contrastada con el gráfico de densidad de una variable aleatoria con distribución normal con parámetros estimados a partir de la muestra³.

De la Figura 1 se puede ver la diferencia en la dispersión de los niveles de ACTH en sangre en pacientes enfermos en relación a los sanos. Además, el contraste con la campana de Gauss nos da indicios de que estamos ante una muestra de una población normal (en el caso de los sujetos sanos), pero no en el caso de la población de enfermos de depresión.

³ Los estimadores utilizados son los mismos que en el caso del test de hipótesis. Obtenidos a partir del método de máxima verosimilitud. Los valores figuran en la Tabla 1.

Niveles de Cortisol en pacientes con depresión y en pacientes sanos.

De forma similar a lo realizado con los niveles de ACTH en sangre, analizamos la relación entre el cortisol en pacientes sanos y enfermos. Los resultados obtenidos se muestran en la Tabla 2.

CORTISOL ESTADO DEL PACIENTE	Media	Desvío	Mediana	MAD	n
Enfermo	261,42	112,24	245,55	102,26	12
Sano	220,30	75,05	212,44	89,98	13

Tabla 2: Medidas de resumen de la variable cortisol, separadas por pacientes enfermos y sanos.

A partir de la Tabla 2 podemos afirmar que los pacientes con trastornos depresivos tienen, en promedio, un mayor nivel de cortisol en sangre en relación a los sujetos sanos; además, el desvío da cuenta de que la dispersión en estos últimos es mucho menor que en aquellos que padecen la enfermedad. Si, a su vez, comparamos la $(MAD/0,6745)=151,6$ con el $Desvío=112,24$ en pacientes enfermos y la $(MAD/0,6745)=133,4$ con el $Desvío=75,05$ en sanos, podemos ver claramente que ambos casos, en promedio, los niveles de cortisol se concentran más alrededor de la media que de la mediana. Luego continuamos nuestro análisis a partir de los siguientes histogramas de la Figura 2.

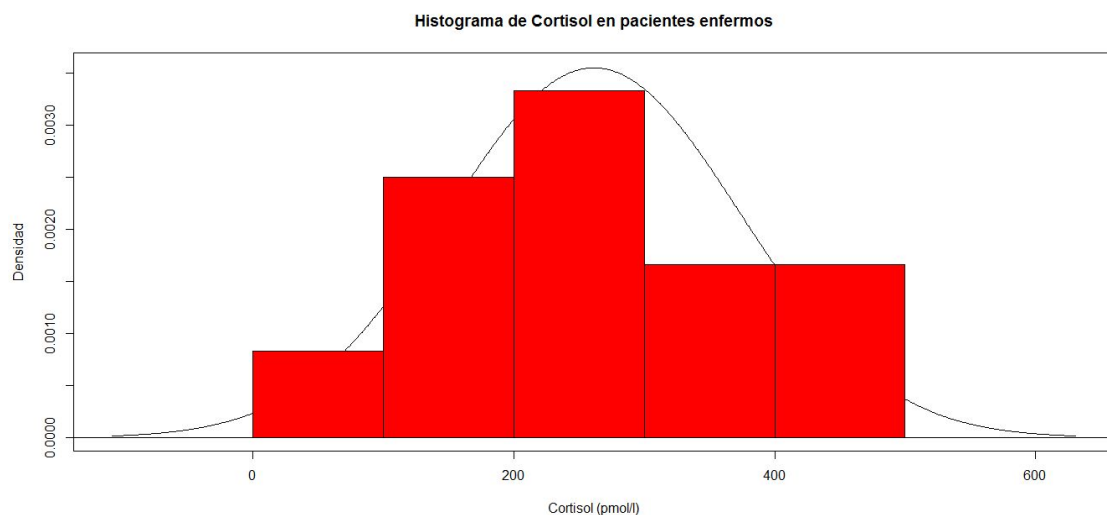


Figura 2: Histogramas de los niveles de cortisol (medidos en pmol/l) en 13 sujetos sanos y 12 enfermos (con depresión unipolar o bipolar) contrastada con el gráfico de densidad de una variable aleatoria con distribución normal con parámetros estimados a partir de la muestra⁴.

De la Figura 2, también se ve como los niveles de cortisol en sangre de los pacientes con trastornos depresivos están mucho más dispersos que los niveles de los sanos. El contraste con la campana de Gauss en el caso de los enfermos, nos da indicios de que podría tratarse de una muestra proveniente de una población normal, ya que se ven los valores centrales con mayor frecuencia y una simetría en el resto de los valores, exceptuando lo que ocurre en la clase de 350 a 400 donde no se observan

⁴ Los estimadores utilizados son los mismos que en el caso del test de hipótesis. Obtenidos a partir del método de máxima verosimilitud. Los valores figuran en la Tabla 2.

datos. En el caso de los sanos, salvo por uno de los datos que resultó alto en relación al resto, se ven agrupados alrededor de la media, lo que daría indicios de una muestra de una población normal. Para analizar la normalidad de las variables cortisol y ACTH que notamos en alguno de los histogramas (Figura 1 y Figura 2), fue que decidimos realizar los gráficos de comparación de cuantiles (QQ-plot) que se muestran en las Figuras 3 y 4.

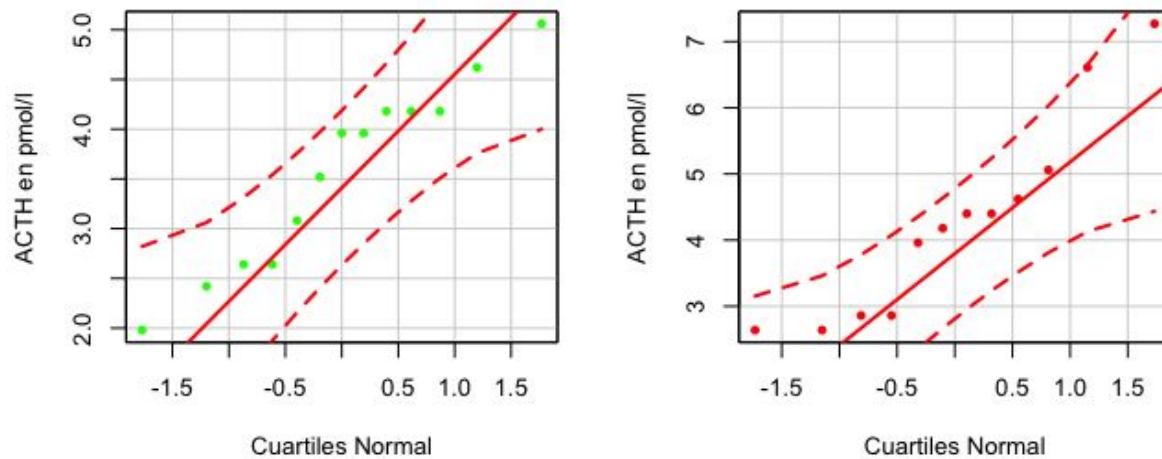


Figura 3: Gráfico de comparación de cuantiles de los niveles de ACTH en sangre en pacientes sanos (izquierda) y enfermos (derecha).

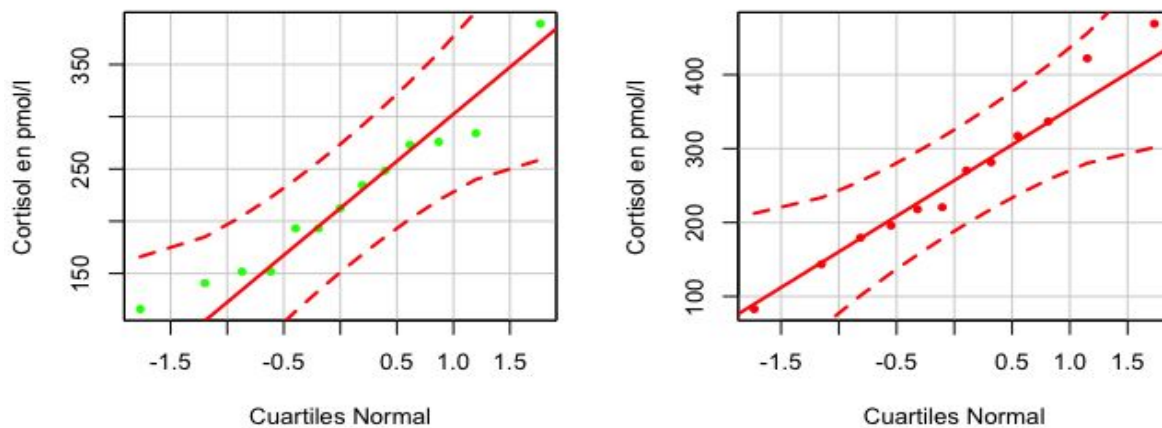


Figura 4: Gráfico de comparación de cuantiles de los niveles de cortisol en sangre en pacientes sanos (izquierda) y enfermos (derecha).

De la Figura 3 y de la Figura 4 podemos afirmar que las variables aleatorias que registran cortisol y ACTH en el grupo de pacientes enfermos y en los sanos se ajustan a una distribución normal.

Como mencionamos anteriormente a partir de la Tabla 1, existen diferencias en los niveles de ACTH en los 12 pacientes con depresión en relación a los 13 sanos. Después de esto determinamos si estas diferencias son significativas como para asegurar que pasará lo mismo en las poblaciones. Para ello, realizamos un test de hipótesis para diferencia de medias de poblaciones normales, teniendo en cuenta que las poblaciones son independientes.

X = niveles de ACTH en sangre (medidos en pmol/l) en pacientes con depresión. $X \sim N(\mu_1; \sigma_1^2)$

Y = niveles de ACTH en sangre (medidos en pmol/l) en pacientes sanos. $Y \sim N(\mu_2; \sigma_2^2)$

$X_1, X_2, \dots, X_{12}$ m.a.s. de $X \sim N(\mu_1; \sigma_1^2)$

$Y_1, Y_2, \dots, Y_{13}$ m.a.s. de $Y \sim N(\mu_2; \sigma_2^2)$

Estimamos los parámetros con el método de máxima verosimilitud. Realizamos una corrección del estimador de la varianza a partir del obtenido con el método ante mencionado para que éste resulte insesgado. De esta forma obtuvimos que $\hat{\mu}_1 = \bar{X}$, $\hat{\mu}_2 = \bar{Y}$, $\sigma_1^2 = s_x^2$ y $\sigma_2^2 = s_y^2$. Estos estimadores son insesgados y de mínima varianza.

$$\bar{x} = 4,29 \quad s_x = 1,5 \quad \bar{y} = 3,57 \quad s_y = 0,94$$

$H_0: \mu_1 - \mu_2 = 0$. La hipótesis nula indica que no hay diferencias entre los niveles de ACTH en sangre en los pacientes depresivos en relación a los sanos.

$H_a: \mu_1 - \mu_2 \neq 0$. Nuestra hipótesis alternativa menciona que existen diferencias en los niveles de ACTH en sangre en los pacientes depresivos en relación a los sanos.

El estadístico de prueba que usamos dependía de si las varianzas desconocidas eran iguales o distintas. Para determinar esto realizamos el siguiente test de hipótesis.

$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1$ Las varianzas son iguales siendo $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$.

$H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1$ Las varianzas son distintas.

El estadístico de prueba es $T = \frac{\frac{s_x^2}{\sigma_1^2}}{\frac{s_y^2}{\sigma_2^2}} \sim F(11; 12)$

$$\text{valor } p = P\left(\frac{s_x^2}{s_y^2} > 2,55 \mid \frac{s_x^2}{s_y^2} \sim F(11; 12)\right) = 0,061$$

A pesar de no resultar un valor p significativamente chico o grande, realizamos el test de hipótesis bajo el supuesto de que las varianzas son distintas porque en estudios relacionados a la medicina se consensúa utilizar un error de 0,05 (Candia y Caiozzi, 2005).

El estadístico de prueba es $T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \sim t(\eta)$

Valor p :

$$P\left(\bar{X} - \bar{Y} > 0,72 \mid \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \sim t(\eta)\right) + P\left(\bar{X} - \bar{Y} < -0,72 \mid \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \sim t(\eta)\right)$$

$$P\left[\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} > \frac{0,72}{\sqrt{\frac{1,5^2}{12} + \frac{0,94^2}{13}}}\right] + P\left[\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} < \frac{-0,72}{\sqrt{\frac{1,5^2}{12} + \frac{0,94^2}{13}}}\right]$$

$$\text{Valor } p = P(t(18) > 1,42) + P(t(18) < -1,42) = 0,086 + 0,086 = 0,17.$$

Analizando la Tabla 2 vimos que existen diferencias en los niveles de cortisol en sangre en los 12 pacientes enfermos en relación a los 13 sanos. Después, determinamos si esas diferencias eran significativas. Para ello realizamos un test de hipótesis para diferencia de medias de poblaciones normales, teniendo en cuenta que son poblaciones independientes.

V = niveles de cortisol en sangre (medidos en pmol/l) en pacientes con depresión. $V \sim N(\mu_3; \sigma_3^2)$

W = niveles de cortisol en sangre (medidos en pmol/l) en pacientes sanos. $W \sim N(\mu_4; \sigma_4^2)$

$$V_1; V_2, \dots; V_{12} \text{ m.a.s. de } V \sim N(\mu_3; \sigma_3^2)$$

$$W_1; W_2, \dots; W_{13} \text{ m.a.s. de } W \sim N(\mu_4; \sigma_4^2)$$

Estimamos los parámetros con el método de máxima verosimilitud. De esta forma obtuvimos que $\hat{\mu}_2 = \bar{V}$, $\hat{\mu}_4 = \bar{W}$, $\sigma_3^2 = s_v^2$ y $\sigma_4^2 = s_w^2$. Estos estimadores son insesgados y de mínima varianza.

$$\bar{v} = 261,42 \quad s_v = 112,24 \quad \bar{w} = 220,3 \quad s_w = 75,05$$

$H_0: \mu_3 - \mu_4 = 0$. La hipótesis nula indica que no hay diferencias entre los niveles de cortisol en sangre en los pacientes depresivos en relación a los sanos.

$H_a: \mu_3 - \mu_4 \neq 0$. Nuestra hipótesis alternativa menciona que existen diferencias en los niveles de cortisol en sangre en los pacientes depresivos en relación a los sanos.

El estadístico de prueba que usamos dependía de si las varianzas desconocidas eran iguales o distintas. Para determinar esto realizamos el siguiente test de hipótesis.

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} \leq 1 \text{ Las varianzas son iguales siendo } \hat{\sigma}_1^2 > \hat{\sigma}_2^2.$$

$$H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1 \text{ Las varianzas son distintas.}$$

$$\text{El estadístico de prueba es } T = \frac{\frac{s_v^2}{\sigma_3^2}}{\frac{s_w^2}{\sigma_4^2}} \sim F(11; 12)$$

$$\text{Valor p} = P\left(\frac{s_v^2}{s_w^2} > 2,2366 \mid \frac{s_v^2}{s_w^2} \sim F(11; 12)\right) = 0,01765$$

Como el valor de p es pequeño vamos a rechazar H_0 , es decir, hay evidencia suficiente para suponer que las varianzas son distintas.

$$\text{El estadístico de prueba es } T = \frac{(\bar{V} - \bar{W}) - (\mu_3 - \mu_4)}{\sqrt{\frac{s_v^2}{n} + \frac{s_w^2}{m}}} \sim t(\eta)$$

Valor p:

$$P\left(\bar{V} - \bar{W} > 41,12 \mid \frac{(\bar{V} - \bar{W}) - (\mu_3 - \mu_4)}{\sqrt{\frac{s_v^2}{n} + \frac{s_w^2}{m}}} \sim t(\eta)\right) + P\left(\bar{V} - \bar{W} < -41,12 \mid \frac{(\bar{V} - \bar{W}) - (\mu_3 - \mu_4)}{\sqrt{\frac{s_v^2}{n} + \frac{s_w^2}{m}}} \sim t(\eta)\right)$$

$$P\left[\frac{(\bar{V} - \bar{W}) - (\mu_3 - \mu_4)}{\sqrt{\frac{s_v^2}{n} + \frac{s_w^2}{m}}} > \frac{41,12}{\sqrt{\frac{112,24^2}{12} + \frac{75,05^2}{13}}}\right] + P\left[\frac{(\bar{V} - \bar{W}) - (\mu_3 - \mu_4)}{\sqrt{\frac{s_v^2}{n} + \frac{s_w^2}{m}}} < \frac{-41,12}{\sqrt{\frac{112,24^2}{12} + \frac{75,05^2}{13}}}\right]$$

$$\text{Valor p} = P(t(19) > 1,067) + P(t(19) < -1,067) = 0,148 + 0,148 = 0,3.$$

Discusión

En la revista SciencyDaily (2014) mencionan una investigación de Wellcome Trust, en donde afirman que síntomas de la depresión, combinados con un alto nivel de cortisol en sangre son suficientes para diagnosticar la enfermedad. En nuestro caso, a partir del análisis realizado sobre las muestras: tablas de estadísticos de resumen (Tabla 1 y Tabla 2) y los histogramas (Figura 1 y Figura 2) observamos que esta relación entre los niveles de cortisol y la enfermedad se evidenciaban en dichas muestras. Si bien las diferencias parecen ser significativas en este caso y también en los niveles de ACTH en sangre en las muestras estudiadas, esas diferencias en ambos casos no resultan significativas para asegurar que lo mismo ocurre en toda la población de pacientes con depresión en relación a aquellos sujetos sanos. Esto nos lleva a pensar que los datos provenientes de la muestra no resultan significativos debido a que la muestra es pequeña.

Conclusiones

En relación a los niveles de cortisol en sangre: a partir de la muestra, encontramos que los niveles de cortisol en sangre en pacientes con trastornos depresivos en comparación con los sujetos sanos resultan mayores. A pesar de esto, no estamos en condiciones de afirmar que los niveles de cortisol en pacientes con trastornos depresivos unipolares y bipolares son distintos que los de los sujetos sanos.

En relación a los niveles de la hormona ACTH en sangre: a partir de la muestra, determinamos un mayor nivel de esta hormona en pacientes que padecen trastornos depresivos en relación a aquellos que no. Sin embargo, esta diferencia no resultó significativa para asegurar que esto sucede en toda la población.

Referencias

- Candia B.R., & Caiozzi A.G. (2005). Intervalos de Confianza. *Revista médica de Chile*, 133(9), 1111-1115. Disponible en: <https://dx.doi.org/10.4067/S0034-98872005000900017> . [20 de noviembre de 2016].
- Curtis, H., Barnes, S., Schnek, A., & Massarini, A. (2008). *Biología* (7ª Edición). Médica Panamericana. Madrid, España.
- Jubiz, W. (1993). Síndrome de Cushing. *Med. UIS*, 7(4), 209-18. Disponible en: <http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&src=google&base=LILACS&lang=p&nextAction=lnk&exprSearch=232166&indexSearch=ID>. [20 de noviembre de 2016].
- Manual diagnóstico y estadístico de los trastornos mentales (15 de noviembre 2016). *Wikipedia, La enciclopedia libre*. Disponible en: https://es.wikipedia.org/w/index.php?title=Manual_diagn%C3%B3stico_y_estad%C3%ADstico_de_los_trastornos_mentales&oldid=95015155. [30 de noviembre de 2016].
- Microsoft Office Excel (2007). (Versión 12.0.6727.5000) SP3. Disponible en: <https://products.office.com/es-es/excel>. [software].
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. (Versión 3.2.2), Vienna, Austria. [software]. Disponible en: <https://www.rproject.org/> . [15 de noviembre de 2016]
- Sachar, E.J., Hellman, L., Roffwarg, H.P., Halpern, F.S., Fukushima, D.K., & Gallagher, T.F. (1973). Disrupted 24-hour patterns of cortisol secretion in psychotic depression. *Archives of General Psychiatry*, 28(1), 19-24.
- Wellcome Trust (17 de febrero 2014). *First biological marker for major depression could enable better diagnosis, treatment*. ScienceDaily. Disponible en: <https://www.sciencedaily.com/releases/2014/02/140217161108.htm>. [15 de noviembre de 2016]
- Wodarz, N., Rupprecht, R., Kornhuber, J., Schmitz, B., Wild, K., Braner, H U., & Riederer, P. (nd). Normal Lymphocyte Responsiveness to Leetins but Impaired Sensitivity to in Vitro Glucocorticoids in Major Depression. *Journal of Affective Disorders*, 22, 241-248.

COMPARACIÓN DE PROPORCIÓN DE FUMADORES PASIVOS Y PERSONAS NO EXPUESTAS AL HUMO DEL TABACO SEGÚN SUS NIVELES DE COTININA

Mariela Weis - Cielo Urquiza

marielaweis92@gmail.com - cielourquiza@gmail.com

Resumen

La cotinina es el principal metabolito de la nicotina, y por lo tanto, el mejor biomarcador de consumo y exposición ambiental al humo de tabaco. En este trabajo se presenta un análisis de la presencia de este alcaloide en el organismo de dos poblaciones: fumadores pasivos y personas no expuestas al humo de tabaco ambiental. Se realizó un estudio descriptivo e inferencial para determinar que los niveles de cotinina superiores a 3,1 ng/ml (valor del límite de detección que permiten diferenciar entre los fumadores y los fumadores pasivos o activos) son más frecuentes en el primer grupo de estudio.

Introducción

Un fumador pasivo es la persona que no fuma pero que se encuentra cerca de aquella que sí lo hace, de manera que ingiere el humo del tabaco aún sin quererlo.

Según el sitio web Salud180 (2013) el humo que inhalan los fumadores pasivos contiene hasta tres veces más nicotina y alquitrán que el que aspira el fumador, y unas cinco veces más monóxido de carbono. Se considera que los fumadores pasivos que están expuestos al humo de tabaco durante una hora, inhalan una cantidad equivalente a dos o tres cigarrillos. Los expertos de la Agencia de Protección Medioambiental (EPA) de los Estados Unidos dividen de tres maneras las principales afecciones o consecuencias de los fumadores pasivos, asociadas con inhalar humo secundario del cigarrillo. Podemos tener consecuencias físicas a corto plazo: tos, problemas dentales (daños a encías, esmalte y piezas dentales), mareos, dolor de cabeza y de garganta, e irritación de los ojos; a mediano plazo: hipertensión arterial, arteriosclerosis, Enfermedad Pulmonar Obstructiva Crónica (EPOC), enfisema pulmonar, asma, ataque cardíaco, infertilidad y daños en el feto (bajo peso, enfermedades genéticas y muerte súbita); y a largo plazo: cáncer (pulmonar, de mama, de garganta, de lengua, de boca, de esófago, de estómago, entre otros).

Se considera que aquellas personas que se encuentran en un ambiente con humo de tabaco, se exponen a dos tipos de ingesta de toxinas como se muestra en la Figura 1: por la corriente principal, siendo ésta la más peligrosa, sucede cuando se inhala el humo que no pasó por el filtro y que contiene una mezcla nociva de gases y partículas de alquitrán carcinogénicas que aportan aproximadamente el 85% de la contaminación atmosférica en interiores; y por la corriente secundaria, aquel humo inhalado, filtrado y exhalado por el fumador. Los no fumadores, que están expuestos al humo del tabaco, se consideran que presentan los mismos riesgos que aquellos catalogados como fumadores leves de cigarrillos. (Tovar Guzmán y López Antuñano, 2000)



Figura 1: Diferentes corrientes de humo del cigarrillo.

Para poder determinar cómo esta exposición afecta al organismo de los fumadores pasivos se necesita realizar algún tipo de medición. Un grupo de profesionales realizó un estudio según el cual “la cotinina es el principal metabolito de la nicotina y es corrientemente considerada el mejor biomarcador de consumo de tabaco en fumadores activos y de exposición ambiental al tabaco en los no fumadores porque tiene una vida media de 18 a 20 horas, mientras que la nicotina tiene una vida media de 1 a 2 horas. Por otro lado, la cotinina puede ser detectada en distintos fluidos biológicos.” (Vachinno, Velurtas, Salinas y Garcialoredo, 2006). En su informe, estos investigadores presentaron sus conclusiones sobre los límites de detección que permiten diferenciar entre los no fumadores y los fumadores pasivos o activos, fijándolo en 3,1 ng/ml (nanogramo por mililitro) de cotinina.

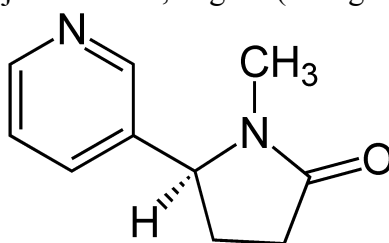


Figura 2: Estructura química de la cotinina.

Por otro lado, en este trabajo nos enfocaremos en comparar la proporción de fumadores pasivos con la de personas no expuestas al humo de tabaco que presentan niveles de cotinina mayores a 3,1 ng/ml en su organismo. Con los resultados que hallaremos deseamos poder mostrar si realmente las personas no expuestas a humo de tabaco tienen el cuerpo libre de cotinina y también poder concientizar a los fumadores pasivos de los peligros a los que están expuestos por vivir en ambientes con humo de tabaco.

Metodología

Los datos utilizados para abordar este trabajo fueron recolectados en la Encuesta Nacional de Salud y Nutrición (NHANES) de Estados Unidos entre los años 1988 y 1994. NHANES es un programa de estudios diseñados para evaluar el estado de salud y nutricional, de adultos y niños en ese país. (CDC, 2016).

Para este trabajo se observaron los niveles de cotinina sérica (en ng/ml) en fumadores activos, en sujetos que no fuman pero están expuestos al humo de tabaco ambiental (fumadores pasivos o NFHTA), y en sujetos que no fuman e indican que no se exponen al humo de tabaco ambiental (NFSHTA).

En base a esas poblaciones se seleccionaron tres muestras aleatorias simples independientes: 40 individuos FUMADORES, 40 sujetos NFHTA y 40 individuos NFSHTA. En todos los casos se registró la cantidad de cotinina (en ng/ml) que se hallaba en su organismo. En particular en este trabajo sólo nos centraremos en observar los niveles de cotinina de las dos últimas muestras.

Se muestran los resultados obtenidos a través del análisis descriptivo e inferencial (test de hipótesis de nivel asintótico para la diferencia de proporciones de las dos muestras trabajadas) realizado sobre las dos poblaciones de estudio, mediante la utilización del software R (versión 3.3.0) y su interfaz gráfica RCommander (versión 2.2-5).

Resultados

Teniendo en cuenta el conjunto de datos utilizado, clasificamos los niveles de cotinina en dos clases, por un lado los niveles superiores a 3,1 ng/ml y por otro lado los inferiores. Luego definimos las variables de estudio correspondientes a cada población como,

X = presencia de cotinina superior a 3,1 ng/ml en un fumador pasivo.

Y = presencia de cotinina superior a 3,1 ng/ml en una persona no expuesta al humo de tabaco.

Estas variables tienen las siguientes distribuciones:

$X \sim Bi(1, p_X)$ siendo $X_1, X_2, \dots, X_{40}$ una m.a.s. de X

$Y \sim Bi(1, p_Y)$ siendo $Y_1, Y_2, \dots, Y_{40}$ una m.a.s. de Y

Con los resultados obtenidos respecto de la medición de niveles de cotinina correspondientes a cada grupo, se realizaron los cálculos relacionados a la estadística descriptiva. Entre ellos, histogramas por grupo, gráficos de barras por grupos donde se muestra la cantidad de personas con niveles de cotinina mayores a 3,1 ng/ml y medidas de resumen necesarias para el análisis.

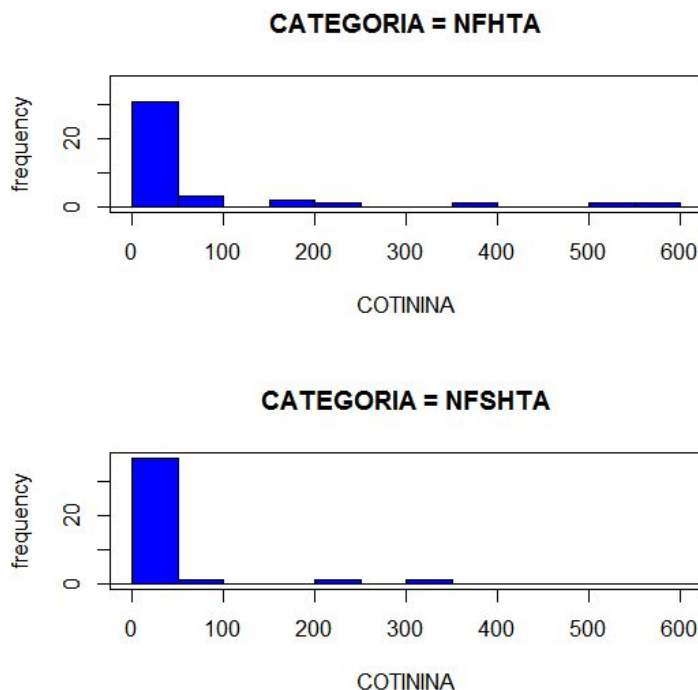


Figura 3: Histogramas por grupos de la cantidad de cotinina presente en el organismo de 40 sujetos NFHTA y 40 NFSHTA

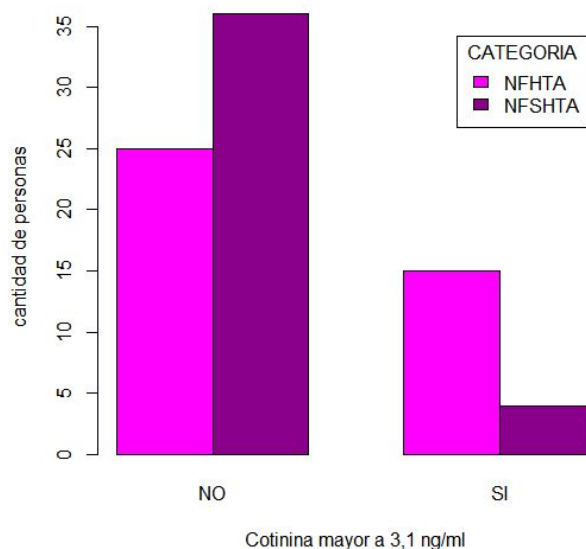


Figura 4: Gráfico de barras comparando presencia de niveles de cotinina superiores a 3,1 entre 40 sujetos NFHTA y 40 NFSHTA.

A continuación, en la Tabla 1, presentamos las medidas de tendencia central y de dispersión correspondientes a cada muestra analizada.

MEDIDAS	COTININA EN NFHTA (ng/ml)	COTININA EN NFSHTA (ng/ml)
<i>Media</i>	60,575	16,350
<i>Desvío</i>	138,084	62,534
<i>Rango intercuartílico</i>	24,5	0
<i>Mínimo</i>	0	0
<i>Máximo</i>	551	309
<i>Asimetría</i>	2,750	4,095
<i>CV (coeficiente de variación)</i>	2,280	3,825
<i>MAD⁵</i>	87,620	29,698
<i>Curtosis</i>	7,030	16,416
<i>Cuantil 25%</i>	1	0
<i>Cuantil 50% (mediana)</i>	1,5	0

⁵ Calculada con <https://www.easycalculation.com/es/statistics/mean-absolute-deviation.php>

Cuantil 75%	25,5	0
-------------	------	---

Tabla 1: medidas de tendencia central y de dispersión de cotinina en NFHTA y en NFSHTA.

Por la información obtenida en base a otras investigaciones sobre este tema, podemos suponer que los fumadores pasivos presentan niveles de cotinina en su organismo en concentraciones mayores a las que presentan los sujetos que no están expuestos al humo de tabaco. Es por eso que queremos demostrar que ésto se cumple también en la población en cuestión en este trabajo. Por lo tanto, decimos que la proporción de fumadores pasivos que presentan niveles de cotinina superiores a 3,1 ng/ml es mayor que la proporción de personas no expuestas al humo del tabaco que tienen cotinina superior a 3,1 ng/ml.

Para comprobar ésto utilizamos un Test de Hipótesis (TH) de nivel asintótico para diferencia de proporciones con un nivel de significación $\alpha = 0,05$.

Pusimos entonces a prueba las siguientes hipótesis:

$$H_0 : p_X \leq p_Y \quad \text{vs} \quad H_a : p_X > p_Y$$

Tomamos el pivote $T = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{m}}}$ donde $\hat{p}_X$ y $\hat{p}_Y$ son los valores estimados de los parámetros de las variables X e Y respectivamente. Podemos decir además que T tiende en distribución a una normal estándar por el Teorema Central del Límite y el Teorema de Slutsky (bajo la hipótesis nula). Obteniendo un valor $p=0,00114$, siendo éste $p \leq \alpha=0,05$ podemos concluir que se rechaza la hipótesis nula (H_0) al nivel del 5%. Además obtuvimos la región de rechazo $RR : \{\hat{p}_X - \hat{p}_Y > -0,148\}$ y como sabemos que $\hat{p}_X - \hat{p}_Y = 0,275 \in RR$ entonces tenemos evidencia suficiente para rechazar a H_0 .

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

$$IC(p_X - p_Y : 95\%) = [0,098; 0,275]$$

Discusión

En la Figura 4 podemos observar claramente que menos de la mitad de los fumadores pasivos presentan niveles de cotinina superiores a 3,1 ng/ml. Sin embargo, este grupo es mayor que el de las personas no expuestas al humo de tabaco con esos valores de cotinina.

Por otro lado, mirando los histogramas de la Figura 3, observamos que en los fumadores pasivos, los niveles de cotinina son más dispersos porque aunque son pocos, igual se hallan concentraciones altas. Sin embargo, en las personas no expuestas al humo de tabaco podemos ver que los valores se encuentran concentrados cercanos a cero.

La asimetría de los niveles de cotinina hallados obliga a utilizar como medida de tendencia central la mediana, siendo la MAD la medida de dispersión asociada. Estos valores se pueden ver en la Tabla 1. Por lo tanto los niveles de cotinina en las muestra de NFHTA y NFSHTA resultan $(1,5 \text{ ng/ml} \pm 87,620 \text{ ng/ml})$ y $(0 \text{ ng/ml} \pm 29.698 \text{ ng/ml})$, respectivamente. Podemos decir con estos valores que la cantidad media de cotinina hallada en los organismos de los NFHTA es más alta y la distribución es más dispersa.

Del TH de nivel asintótico para diferencia de proporciones que realizamos podemos afirmar, con un nivel de significación del 5%, que que hay mayor proporción de fumadores pasivos con niveles de cotinina mayores a 3,1 ng/ml que personas no expuestas al humo de tabaco con estos niveles.

Además el IC hallado nos permite encontrar valores factibles para nuestro parámetro de interés, pudiendo concluir que la diferencia entre las proporciones de NFHTA y NFSHTA que presentan valores de cotinina mayores a 3,1 ng/ml ronda entre 0,098 y 0,275 ng/ml, con una probabilidad del 95%. Por lo tanto como el IC construido no contiene al cero, se ratifica nuestra conclusión anterior.

Conclusiones

El objetivo de estudio de este trabajo consistió en comparar la proporción de fumadores pasivos con niveles de cotinina mayores a 3,1 ng/ml con la proporción de personas no expuestas al humo de tabaco que también superen estas cantidades de cotinina. Luego de los distintos análisis llevados a cabo concluimos, mediante inferencia estadística con un nivel de significación del 5%, que la proporción de fumadores pasivos que presentan niveles de cotinina superiores a 3,1 ng/ml es mayor que la proporción de personas no expuestas a humo de tabaco con estos mismos niveles. En resumen, aunque los fumadores pasivos no son consumidores activos, igual presentan en su organismo rastros del humo de tabaco que son significativos, y pueden derivar en algunas enfermedades. Sin embargo, las personas no expuestas al humo de tabaco casi no registran niveles de cotinina superiores a 3,1 lo cual nos dice que tienen menor probabilidad de contraer alguna enfermedad de las antes mencionadas.

Referencias

- CDC (24 de febrero de 2016), *National Health and Nutrition Examination Survey*. [en línea]. Disponible en: <http://www.cdc.gov/nchs/nhanes/index.htm>
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [software]. Disponible en: <https://www.r-project.org/>
- Salud180 (2013). *10 efectos del cigarrillo en fumadores pasivos*. [en línea]. Disponible en: <http://www.salud180.com/jovenes/10-efectos-del-cigarrillo-en-fumadores-pasivos>
- Tovar Guzmán, V.J.; López Antuñano, F.J. (2000). *Fumadores involuntarios: Exposición pasiva al humo de tabaco en el ambiente*. Revista del Instituto Nacional de Enfermedades Respiratorias de México, 13 (4): pp. 233-239
- Vacchino, M.N.; Velurtas, S.M.; Salinas, G.P.; Garcialoredo, H.H. (2006). *Determinación de cotinina y exposición al tabaco*. Federación Bioquímica de la Provincia de Buenos Aires. Acta bioquímica clínica latinoamericana, 40 (2) [en línea]. Disponible en: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0325-29572006000200004

HIDROCEFALIA Y CIRCUNFERENCIA DE LA CABEZA

Dalma Cañada - Yohana Esnal - Soledad Felipe

dalma.c23@gmail.com - yohana.esnal.1991@gmail.com - solepedrito@gmail.com

Resumen

La hidrocefalia afecta a uno de cada 500 niños que nacen. Es causada por una amplia variedad de problemas médicos, y las circunstancias de la enfermedad de cada niño son únicas. Uno de los métodos para detectar la hidrocefalia es mediante la medición de la circunferencia de la cabeza del bebé. En los últimos años, han habido notables adelantos en el tratamiento de la hidrocefalia, por eso es muy importante su pronto diagnóstico.

En este reporte se pretende analizar los datos correspondientes a la circunferencia de la cabeza de niños y niñas de dos meses de edad, para determinar la posibilidad de encontrarnos con este tipo de enfermedad. Asimismo, estudiaremos por qué nos encontramos en la vida cotidiana con que hay distintas tablas de mediciones de acuerdo al sexo del niño@.

Introducción

Como encontramos en la página de MedlinePlus (2016), la hidrocefalia es un trastorno cuya principal característica es la acumulación excesiva de líquido cefalorraquídeo (LCR) en el cerebro. Los niveles de LCR pueden elevarse debido a varias causas, una de ellas es si se bloquea el flujo del líquido; otra es que éste no logre absorberse apropiadamente en la sangre o que el cerebro produzca demasiada cantidad de dicho líquido. Demasiado LCR ejerce presión sobre el cerebro, lo cual lo empuja hacia arriba contra el cráneo y daña el tejido cerebral.

La hidrocefalia puede empezar mientras el bebé está creciendo en el útero y afecta a uno de cada 500 niños que nacen (National Institute of Neurological Disorders and Stroke, 2010). Es común en bebés que tienen mielomeningocele, un defecto congénito en el cual la columna vertebral no se cierra apropiadamente. También puede deberse a otros defectos congénitos; ciertas infecciones durante el embarazo, sangrado intracerebral durante o poco después del parto (especialmente en bebés prematuros); por una lesión antes, durante o después del parto, que incluye hemorragia subaracnoidea.

En los bebés, la hidrocefalia provoca que la fontanela (zona blanda) aumente su tamaño, por lo tanto la cabeza será más grande de lo esperado. Al momento de examinar a un bebé hay varios detalles que son observados para determinar si éste tiene hidrocefalia, alguna de las observaciones son: si sus venas están hinchadas o distendidas en el cuero cabelludo; si al momento que el médico golpea suavemente el cráneo del niño con la punta de los dedos los sonidos son anormales, lo que sugiere un problema en los huesos craneales; si la cabeza del pequeño es una parte o toda más grande de lo normal; si los ojos del niño lucen “hundidos”; los reflejos pueden ser normales.

Para llevar un control se realizan repetidas mediciones del perímetro cefálico, esto nos muestra si la cabeza se está agrandando. Uno de los mejores exámenes ante la sospecha de hidrocefalia es la tomografía computarizada.

Es muy importante su pronto diagnóstico, ya que sin tratamiento, las personas con hidrocefalia pueden sufrir diferentes grados de discapacidades intelectuales, físicas y neurológicas.

Una vez diagnosticada, el objetivo del tratamiento es reducir o prevenir el daño cerebral a través del mejoramiento del flujo del LCR (Edwards & Derechin, 2009). El tratamiento suele incluir una cirugía para implantar una sonda. Ésta es un tubo flexible pero robusto, que desvía el flujo del LCR

hacia el abdomen en donde puede reabsorberse. Los medicamentos y la rehabilitación también pueden ayudar.

Los siguientes gráficos fueron extraídos de la Libreta Sanitaria Infanto Juvenil, brindada por el Ministerio de Salud de la provincia de La Pampa, que se extrajo y reprodujo de la Organización Mundial de la Salud. En ellos se indica la curva de crecimiento del perímetro cefálico (en cm) de niñas y niños, según los meses-años, respectivamente.

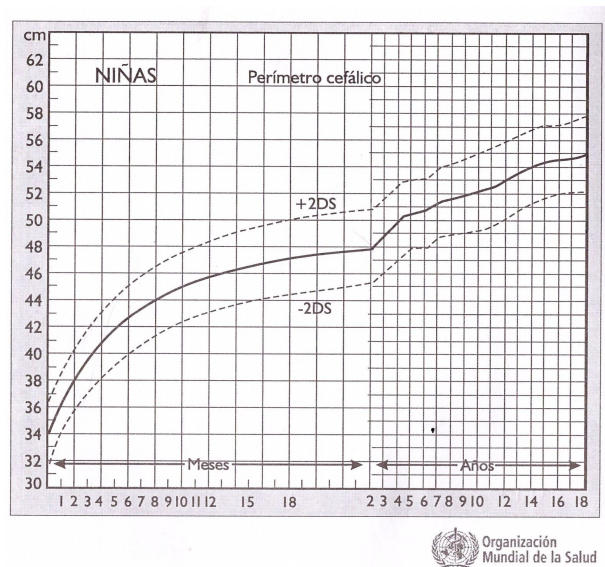


Figura 1: Gráfico del perímetro cefálico de las niñas en la Libreta sanitaria infanto juvenil.

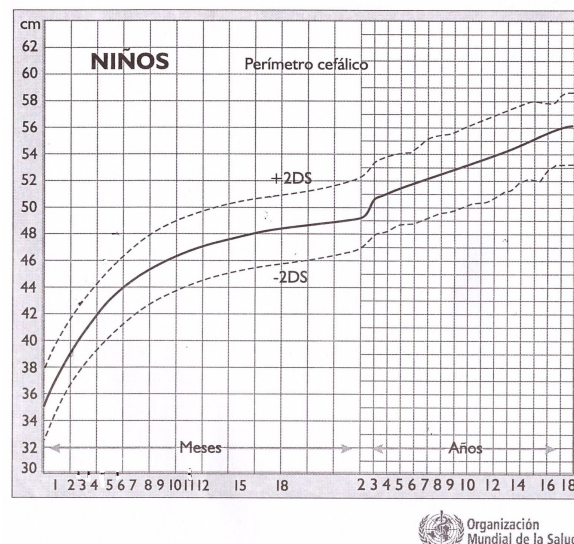


Figura 2: Gráfico del perímetro cefálico de los niños en la Libreta sanitaria infanto juvenil.

En estos gráficos podemos observar cuáles son las medidas estándares del perímetro cefálico de los niños y niñas desde su nacimiento hasta los 18 años de edad.

Bajo estas perspectivas, los objetivos que tenemos con nuestro trabajo son:

- ❖ verificar si la proporción de niñ@s que podrían sufrir hidrocefalia es efectivamente 1 de 500
- ❖ comparar la circunferencia cefálica media de niñas y niños de dos meses de edad.

Metodología

El estudio se realizó sobre una muestra aleatoria simple correspondiente a niños y niñas de dos meses de edad a los cuales se les midió el perímetro de la circunferencia de la cabeza.

Los 100 datos fueron obtenidos de la tercera Encuesta Nacional de Salud y Nutrición (NHANES III), perteneciente al *Department of Health and Human Services* de EE.UU., entre 1988 y 1994.

Se desarrolló un análisis descriptivo sobre cada una de las medidas consideradas y se aplicaron técnicas inferenciales utilizando el software estadístico R (versión i386 3.3.1), a través de su interfaz gráfica R Commander (versión 2.2-5).

Para llevar a cabo el primer objetivo, primero definimos nuestra variable:

W = tendencia a la hidrocefalia. Donde determinamos los límites para clasificar si un niñ@ tiene tendencia a tener hidrocefalia o no y luego registramos qué cantidad de niños teníamos en cada grupo. A continuación, utilizamos un test de hipótesis (TH) para la media de la variable estudiada con un nivel de significancia de 0,01.

Para la realización del segundo objetivo, definimos las dos variables a utilizar:

X = medida de la circunferencia cefálica, en centímetros, de las niñas de dos meses de edad.

Y = medida de la circunferencia cefálica, en centímetros, de los niños de dos meses de edad.

donde primero se puso a prueba la normalidad de las dos variables utilizadas, luego se realizó un test de hipótesis para el cociente de varianzas, con un nivel de 0,01 para determinar si éstas eran iguales o no y finalmente se desarrolló un TH con un nivel de 0,01 para la diferencia de medias de las distintas variables.

Resultados:

A continuación se mostrarán los resultados obtenidos para llevar adelante nuestro primer objetivo.

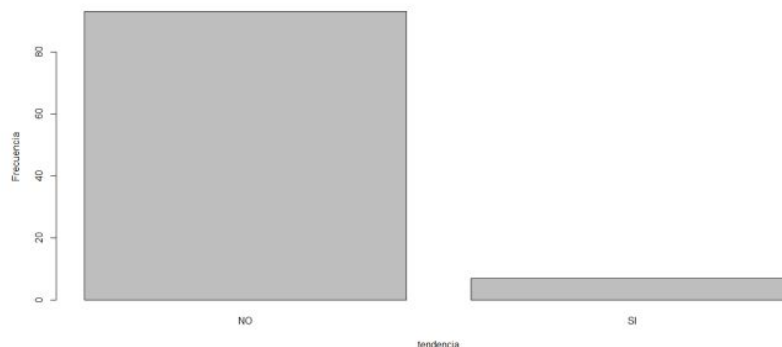


Figura 3: Gráfico de barras de la tendencia (o no) de sufrir hidrocefalia de niñ@s de dos meses de edad.

En la **Figura 3** podemos observar que los niños pueden tener o no tendencia a sufrir hidrocefalia, por lo tanto suponemos que tiene distribución Bernoulli. Lo que observamos allí también es que hay 93 niñ@s que no tienen tendencia y 7 niñ@s que si tienen tendencia a sufrir hidrocefalia.

Para verificar si la proporción de niñ@s que podrían sufrir hidrocefalia es efectivamente 1 de 500, primero analizamos la m.a.s. compuesta por la circunferencia cefálica de 100 niñ@s de dos meses de edad. Consideramos, observando los gráficos de la circunferencia cefálica de la libreta infanto juvenil que, si la circunferencia cefálica de los niños es mayor o igual a 43 cm y la de las niñas es

mayor o igual a 42 cm, entonces tienen tendencia a sufrir hidrocefalia.

De esta manera, podemos afirmar que la variable aleatoria W tiene distribución Bernoulli (ya que tiene o no tendencia a sufrir el trastorno) con parámetro estimado $\hat{p} = 0,07$.

El valor $\hat{p}$ lo estimamos a partir de la muestra dada por el método de máxima verosimilitud, tal como lo demostraron Caffarone et al. (2016), $\hat{p} = \bar{w} = 0,07$. Luego, construimos un TH con nivel asintótico, donde las hipótesis que establecimos fueron: $H_0 : p = \frac{1}{500}$ vs. $H_a : p \neq \frac{1}{500}$

De esta manera, según las Notas de clase. Estadística para futuros matemáticos (Dieser, 2016), el estadístico tiene distribución normal estándar, con $\alpha = 0,01$, valor $p = 0,008 < \alpha = 0,01$, por lo tanto, tenemos suficiente evidencia para rechazar H_0 . El IC correspondiente es $IC(p:99\%) = [0,0042; 0,1358]$, es decir que la proporción de niñ@s que tiene tendencia a sufrir hidrocefalia se encontrará dentro de esos valores.

Para el segundo objetivo, en primer lugar vamos a considerar las dos variables (niños y niñas) por separado:

➤ Para el caso de las niñas, definimos la variable aleatoria:

X = medida de la circunferencia cefálica, en centímetros, de las niñas de dos meses de edad.

La media y el desvío obtenido a partir de la muestra son 40,05 cm y 1,64 cm, respectivamente.

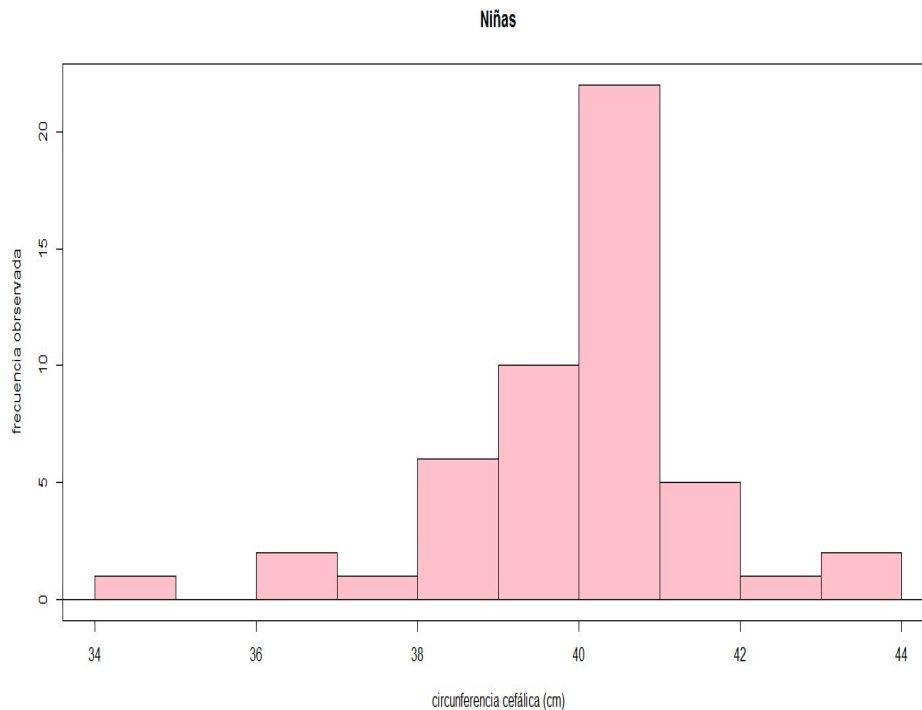


Figura 4: Histograma de frecuencias de la medida de circunferencia cefálica de las niñas, en cm.

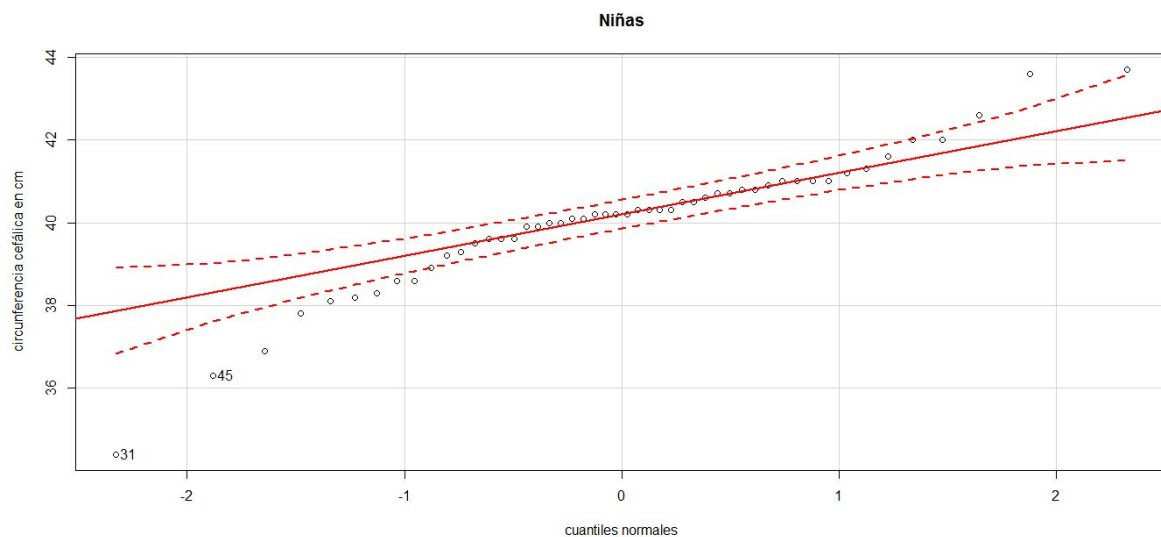


Figura 5: QQ-plot de la medida de la circunferencia cefálica de las niñas, en cm.

En la **Figura 4** podemos observar que hay más de 20 niñas que tienen una medida de la circunferencia cefálica entre 40 y 41 cm, siendo ésta la mayor frecuencia observada. Respecto a las que miden entre 34 a 35 cm, entre 37 a 38 cm y de 42 a 43 cm se observa que tienen la misma frecuencia. La máxima medida de circunferencia está entre 43 cm y 44 cm, que corresponden a 2 niñas, la misma cantidad de niñas que tienen una medida de 36 a 37 cm. No se encontraron niñas que tuvieran una medida de 35 a 36 cm. Encontramos 10 niñas con circunferencia cefálica entre 39 a 40 cm, 5 de 41 a 42 cm, 6 que tienen una circunferencia cefálica de 38 a 39 cm.

En la **Figura 5** en el QQ-plot los puntos se acumulan, en su gran mayoría, alrededor de una recta, lo que nos da un indicio de que los datos provienen de una distribución normal. Dado que los parámetros poblacionales son desconocidos, los estimamos a través de la muestra, por el trabajo realizado por Dieser et al. (2016) sabemos que es IMVU, suficiente, eficiente y consistente. Los parámetros son estimados y resultaron : $\hat{\mu}_1 = \bar{x} = 40,05$ cm y $\hat{\sigma}_1^2 = s_1^2 = 2,69$ cm, para la media y desvío, respectivamente.

Se realizó un test de bondad de ajuste chi-cuadrado para probar la normalidad de la variable, utilizando la media y la varianza estimadas, (chi-cuadrado = 11,345, gl = 3).

➤ Para el caso de los niños, definimos la variable aleatoria:

Y = medida de la circunferencia cefálica, en centímetros, de los niños de dos meses de edad.

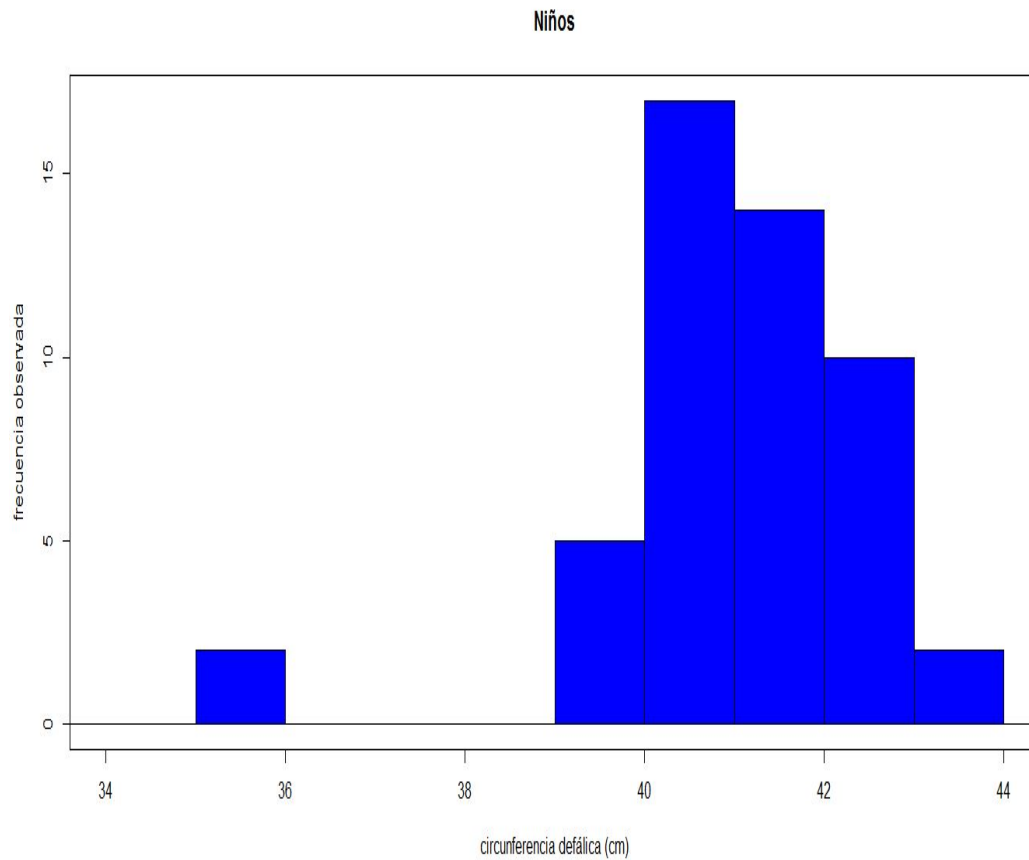


Figura 6: Histograma de la circunferencia cefálica de los niños, en cm.

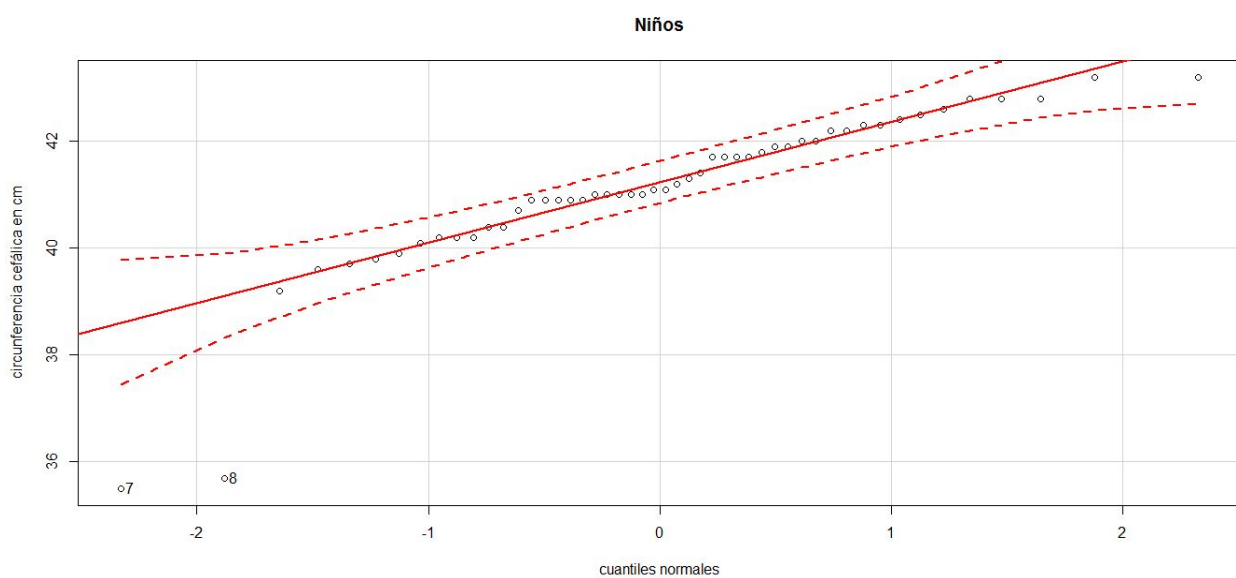


Figura 7: QQ-plot de la medida de la circunferencia cefálica de las niños, en cm.

En la **Figura 6** podemos observar que la mayor cantidad de niños (17) tienen una circunferencia cefálica entre 40 a 41 cm, no se encuentran niños que tengan una circunferencia entre 36 a 39 cm. Se registraron que, de 35 a 36 cm y 43 a 44 cm se encuentran en la misma cantidad (2). De 41 a 42 cm encontramos 14 niños y de 42 a 43 cm, se encuentran 10 niños. Y luego encontramos 5 niños que tienen una circunferencia cefálica de 39 a 40 cm.

En la **Figura 7** en el QQ-plot los puntos se acomodan alrededor de una recta, por lo que podemos afirmar que los datos provienen de una población normal.

Como dijimos anteriormente, estimamos los parámetros poblacionales a partir de la muestra, obteniendo $\hat{\mu}_2 = \bar{y} = 41,10$ cm y $\hat{\sigma}_2^2 = s_2^2 = 2,24$ cm, para la media y desvío, respectivamente.

Se realizó un test de bondad de ajuste chi-cuadrado para probar la normalidad de la variable, utilizando la media y la varianza estimadas, (chi-cuadrado = 11,345, gl = 3).

Como resumen de los valores obtenidos, presentamos la siguiente tabla:

Variable aleatoria	Medida de tendencia central: la media	Medida de dispersión: el desvío
X	40,05	1,64
Y	41,10	1,50

Tabla 1: Medida de tendencia central y dispersión de las variables, medidas en cm.

Más tarde se compararon las circunferencias cefálicas medias de niñas y niños de dos meses de edad. Para ésto fue necesario primero comprobar si las varianzas eran iguales, con un test para el cociente de varianzas de poblaciones normales para muestras aleatorias independientes (F de Fisher (49,49) grados de libertad, valor $p = 0,26 > \alpha = 0,01$), por lo que no se pudo rechazar la hipótesis, luego las varianzas eran iguales. Luego se realizó un TH para la diferencia de medias donde nuestras hipótesis fueron: $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$, donde $\hat{\mu}_1 = 40.05$ y $\hat{\mu}_2 = 41.10$, el estadístico tiene distribución t de Student con 98 grados de libertad, con una probabilidad de 0,01 y el desvío ponderado es 1,57, valor $p = 0,0012 < \alpha = 0,01$, por lo cual podemos rechazar H_0 , es decir las medias son distintas. Por lo tanto se construyó el IC correspondiente, $IC(\mu_1 - \mu_2 : 99\%) = [-1,88; -0,22]$, donde el intervalo no contiene al cero, es decir ratificamos nuestra conclusión anterior. Además podemos afirmar que la diferencia de la circunferencia cefálica media de niños es mayor que la de las niñas de dos meses de edad y esa diferencia ronda los 1,88 y 0,22 unidades, con una probabilidad del 99%.

Discusión

Como respuesta a nuestro primer objetivo, podemos decir que al analizar los datos, nos encontramos con una distribución Bernoulli, donde observamos que el 93% de los niños se encuentran dentro de los valores usuales, con respecto a su circunferencia cefálica (por **Figura 3**).

Luego, mediante un TH con nivel de significación asintótico no podemos afirmar que uno de cada 500 niños tienen tendencia a sufrir hidrocefalia, como encontramos en *National Institute of Neurological Disorders and Stroke* (2010). La proporción de niños que tienen tendencia a sufrir hidrocefalia ronda entre [0,0042; 0,1358].

Para nuestro segundo objetivo, podemos decir que en más del 80% de las medidas de la

circunferencia cefálica de las niñas de dos meses se encuentran entre 38 y 42 cm, dentro de los cuales se puede observar una mayor acumulación entre 40 y 41 cm, tal como lo muestra el gráfico de la **Figura 4**. En el QQ-plot de la **Figura 5** se puede apreciar que la gran mayoría de los puntos se acumulan alrededor de una recta. Por este motivo se supuso que la variable circunferencia cefálica de las niñas de dos meses (en cm), se distribuye normalmente. El test de bondad de ajuste realizado revela que hay evidencia al nivel del 1% para suponer que los datos de la muestra provienen de una distribución normal con media 40,05 cm y desvío 1,64 cm. Luego, podemos resumir la circunferencia cefálica de las 50 niñas de dos meses como $(40,05 \pm 1,64)$ cm.

Algo similar podemos decir sobre la circunferencia cefálica de los niños de dos meses (en cm), donde más del 90% de las medidas se encuentran entre 39 y 43 cm, dentro de los cuales se puede observar una mayor acumulación entre 40 y 41 cm, tal como lo muestra el gráfico de la **Figura 6**. En el QQ-plot de la **Figura 7**, también se puede apreciar una acumulación de los puntos alrededor de una recta. Por este motivo, al igual que con las niñas, se supuso que la variable circunferencia cefálica de los niños de dos meses (en cm), se distribuye normalmente. El test de bondad de ajuste realizado revela que hay evidencia al nivel del 1% para suponer que los datos de la muestra provienen de una distribución normal con media 41,10 cm y desvío 1,5 cm. Luego, podemos resumir a la circunferencia cefálica de los 50 niños de dos meses como $(41,10 \pm 1,5)$ cm.

En cuanto a la relación del estudio de nuestras variables con respecto a las medidas determinadas por el Ministerio de Salud de la provincia, 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 las circunferencias cefálicas de los 50 niños y 50 niñas de Estados Unidos, en media, no son iguales.

Además, con un 99% de confianza, podemos afirmar que las medias de niñas y niños difieren, y además la circunferencia cefálica media de las niñas es menor que la de los niños, ya que el intervalo resultó con signo negativo. Podemos asegurar que en promedio la diferencia de medias es $-1,05 \pm 0,83$.

Conclusiones

Uno de los objetivos de estudio de este proyecto consistió en verificar si la proporción de niñ@s que podrían sufrir hidrocefalia era efectivamente 1 de 500. Luego del análisis realizado llegamos a la conclusión con nuestros datos de que esto no se cumple, y que la tendencia a sufrirla es mayor. Por lo que no podemos afirmar que la información que se brinda en la web es verídica, pero si esto fuera así, podríamos concluir que hay niñ@s que tienen tendencia pero finalmente no sufren este trastorno. El segundo objetivo fue comparar la circunferencia cefálica media de niñas y niños de dos meses de edad. Que después de realizar varios estudios llegamos a la conclusión de que la circunferencia media de niños y niñas difiere, y es por eso que en las libretas sanitarias del Ministerio de Salud aparecen diferenciadas. Y además, podemos afirmar que la circunferencia cefálica media de los niños es mayor que la circunferencia cefálica media de las niñas, aunque la diferencia no es muy amplia, pero en niños tan pequeños puede ser significativo. Por esta razón podemos justificar la existencia de dos cuadros diferenciados en las libretas sanitarias infanto juvenil.

Referencias

- Caffarone, M.L.; Peralta, R. & Weis, M. (2016). *Estimación Puntual: Distribución Bernoulli*. Material no publicado. FCEyN-UNLPam. Santa Rosa-La Pampa [10 de noviembre de 2016]
Dieser, M.P. (2016). *Estadística para futuros matemáticos: Notas de clase*. Material no publicado. FCEyN-UNLPam. Santa Rosa-La Pampa. [4 de noviembre de 2016]

- Dieser, M.P.; Funkner, S. & Caverio, L. (2016). *Estimadores de la media de una población con distribución normal*. Material no publicado. FCEyN-UNLPam. Santa Rosa-La Pampa. [10 de noviembre de 2016]
- Edwards, M. & Derechin, M. (2009). *Sobre la Hidrocefalia* [página web]. San Francisco, California: Susan Eastwood. Disponible en: http://www.hydroassoc.org/docs/Sobre_la_Hidrocefalia_web-09.pdf [4 de noviembre de 2016]
- MedlinePlus. (01 de noviembre de 2016). *Hidrocefalia*. [página web]. Disponible en: <https://medlineplus.gov/spanish/ency/article/001571.htm> [4 de noviembre de 2016]
- National Institute of Neurological Disorders and Stroke. (03 de marzo de 2010). *Hidrocefalia*. [página web]. Disponible en: https://espanol.ninds.nih.gov/trastornos/la_hidrocefalia.htm [7 de noviembre de 2016]
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [software]. Disponible en <https://www.R-project.org/>.