

[Psicología de ayer y hoy]

[10.22402/j.rdipecs.unam.6.2.2020.302.499-517](https://doi.org/10.22402/j.rdipecs.unam.6.2.2020.302.499-517)

ESTADÍSTICA ROBUSTA APLICADA A LAS MEDIDAS DE LOCALIZACIÓN Y ESCALA: NOTA TÉCNICA

David Ruiz Méndez, Mirna Elizabeth Quezada y Cynthia Zaira Vega Valero
Instituto de Estudios Superiores Monterrey
México

RESUMEN

Cuando se hace la medición de una variable es común que los estadígrafos más utilizados, como la media aritmética y su error estándar asociado, sean estimadores imprecisos para describir los datos de una muestra y poder hacer inferencias. Para solucionar esta situación, el investigador puede hacer uso de la estadística robusta. Dicha estadística ofrece una serie de estimadores alternativos resistentes a la influencia de los datos atípicos en una distribución. Su uso resulta en información e inferencias más precisas. El objetivo de este artículo es describir un conjunto de procedimientos para calcular medidas de localización y escala con métodos robustos usando el lenguaje de programación R y el software estadístico SPSS. Primero se revisan métodos de detección de datos atípicos de manera visual y cuantitativa. Después se revisan opciones para las medidas de localización, como la media recortada, la media winsorizada y el estimador M. Cada medida estará acompañada de su correspondiente error estándar. Por último, se presentan medidas de escala, como el rango intercuartil y su modificación denominada cuartos ideales. Se concluye invitando al lector al uso razonado de los procedimientos, en congruencia con sus posibilidades, intereses e implicaciones teóricas y metodológicas.

Palabras Clave:

estadística robusta, medidas robustas de localización, medidas robustas de escala, valores atípicos, distribuciones sesgadas.

ROBUST STATISTICS APPLIED TO LOCATION AND SCALE MEASUREMENTS: TECHNICAL NOTE

ABSTRACT

When measuring psychological variables, it is usual to find that most of the commonly used statistics, such as the arithmetic mean and its associated standard error are, at best, imprecise in the task of data summary and inference. To overcome this situation, researchers can make use of robust statistics. Robust statistics provide a series of alternative estimates resistant to the influence of outliers. Their use results in precise information and inferences. The aim of this paper is to describe a group of procedures to estimate measures of location and scale with robust methods using R and IBM SPSS software. First, the different visual and quantitative methods for detecting outliers are reviewed. Then, different alternatives of location measures are reviewed such as the trimmed mean, the winsorized mean and M estimators. Each location measure will be presented with its associated standard error. Last, some scale measures are presented, such as the interquartile range and a modification, called ideal fourths. Conclusions emphasize the thoughtful use of the procedures in relation to the reader possibilities, interests and the theoretical and methodological implications.

Keywords:

robust statistics, robust location measures, robust scale measures, outliers, skewed distributions.

BITÁCORA DEL ARTÍCULO:

| Recibido: 13 de mayo de 2020 | Aceptado: 12 de junio de 2020 | Publicado en línea: Julio-Diciembre de 2020 |

AUTORÍA Y DERECHOS DE PROPIEDAD INTELECTUAL

ESTADÍSTICA ROBUSTA APLICADA A LAS MEDIDAS DE LOCALIZACIÓN Y ESCALA: NOTA TÉCNICA

David Ruiz Méndez, Mirna Elizabeth Quezada y Cynthia Zaira Vega Valero
Instituto de Estudios Superiores Monterrey
México



David Ruiz Méndez
ITESM
Correo: davidm@tec.mx

Licenciado y maestro en psicología por la UNAM. Ha participado en Congresos especializados internacionales en relación con la evaluación psicológica en las organizaciones y análisis experimental y cuantitativo en psicología. Actualmente estudia el doctorado en Análisis Experimental de la Conducta en la UNAM y se desempeña como profesor en el Tecnológico de Monterrey Campus Estado de México.



Mirna Elizabeth Quezada
FES Iztacala UNAM
Correo: psc.m.quezada@gmail.com

Licenciada en Psicóloga por Universidad Autónoma de Tamaulipas; Maestra en Psicología, Gestión Organizacional por la Facultad de Estudios Superiores Iztacala UNAM, becaria de los programas PAPIME, PAPIIT y CONACyT. Actual docente del programa de Maestría en Psicología de la UNAM, de la Lic. en Psicología del Sistema de Universidad Abierta y Educación a Distancia-UNAM, y docente de la Licenciatura en línea de la Universidad Tecnológica de México.

CONTRIBUCIÓN DE LOS AUTORES

David Ruiz Méndez colaboró en la introducción del artículo, el desarrollo matemático de los estadísticos en la primera sección y la elaboración de los análisis en R. | Mirna Elizabeth Quezada contribuyó en la introducción así como en el desarrollo de los análisis estadísticos en IBSM SPSS. | Cynthia Zaira Vega Valero contribuyó en la elaboración de la introducción y revisión de los procesos estadísticos.

AGRADECIMIENTOS

Este artículo forma parte de una serie de notas técnicas enfocadas a discutir algunos métodos de utilidad de la estadística robusta para el análisis de datos de los científicos de la conducta.
Agradecimiento al proyecto: PAPIIT IT300218.

DATOS DE FILIACIÓN DE LOS AUTORES

Instituto de Estudios Superiores Monterrey | Facultad de Estudios Superiores Iztacala UNAM



Copyright: © 2020 Ruiz-Méndez, D.; Quezada, M.E. & Vega-Valero, C.Z.

Este es un artículo de acceso abierto distribuido bajo los términos de la licencia [Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/), por lo que su contenido gráfico y escrito se puede compartir, copiar y redistribuir total o parcialmente sin necesidad de permiso expreso de sus autores con la única condición de que no se puede usar con fines directamente comerciales y los términos legales de cualquier trabajo derivado deben ser los mismos que se expresan en la presente declaración. La única condición es que se cite la fuente con referencia a la [Revista Digital Internacional de Psicología y Ciencia Social](https://doi.org/10.24018/psicologia-y-ciencia-social) y a sus autores.

TABLA DE CONTENIDO

INTRODUCCIÓN	508
FUNDAMENTOS TEÓRICOS DEL PROCEDIMIENTO	508
DETECCIÓN DE OUTLIERS Y MEDIDAS ROBUSTAS DE LOCALIZACIÓN USANDO R	509
MEDIDAS ROBUSTAS DE ESCALA EMPLEANDO R	511
ESTADÍSTICA ROBUSTA DE LOCALIZACIÓN Y ESCALA EMPLEANDO SPSS COMO SOFTWARE ESTADÍSTICO	512
CONCLUSIÓN	516
REFERENCIAS	517

INTRODUCCIÓN

Cuando un investigador de las ciencias de la conducta hace la medición de alguna variable, el producto final es un conjunto de valores numéricos que tienen cierto grado de isomorfismo con los atributos o dimensiones que se busca medir (Nunnally y Bernstein, 1994). Una vez obtenidos los números, se busca extraer información útil a partir de ellos. El investigador sabe que este proceso lo ayudará a contestar las preguntas que se planteó al principio de su tarea investigativa. La estadística descriptiva es la rama de la estadística que le permite al investigador extraer información relevante de una distribución de datos numéricos (Hopkins & Glass, 1978). El objetivo de esta rama es servir como herramienta para la toma de decisiones con base en la presentación ordenada y descripción de los datos extraídos mediante la observación (Fernández, Sánchez, Córdoba y Largo, 2002).

Desde un punto de vista matemático, la extracción de información de la distribución de datos se puede efectuar a partir de la identificación de sus momentos (Press, Teukolsky, Vetterling y Flannery, 2007). Los momentos, o la suma sucesiva de las potencias de los números enteros de una distribución de datos, engloban un conjunto de medidas cuantitativas que proporcionan información de una distribución. A pesar de la utilidad de las medias producidas en cada momento, es común que, en determinadas condiciones, las medidas utilizadas suelen presentar información distorsionada (Casella y Berger, 2002).

Un indicador importante para calcular el grado de susceptibilidad de una estadística a la distorsión es la estimación de su punto de ruptura de una muestra finita (*fine-sample breakdown point*), el cual hace referencia a la proporción más pequeña de observaciones que, al ser alteradas, hacen que el estadístico pierda su valor informativo (Maronna, Martín, Yohai y Salibán-Barrera, 2019; Wilcox, 2017). El punto finito de ruptura de la muestra para las medidas más comunes suele ser muy bajo y el origen de la distorsión puede provenir de datos denominados "outliers". Los outliers son puntuaciones muy diferentes al resto del conjunto de datos, cuya naturaleza extrema y de alta influencia llega a afectar de modo sustancial el cálculo de los estadígrafos de una distribución y la inferencia de parámetros a partir de ellos (Field, 2013). Una solución a estos problemas es el uso de la estadística robusta.

El término "robusto" se refiere a la resistencia de una estadística a producir error cuando hay una desviación en los supuestos básicos de su cálculo e interpretación (Huber, 1981). Este tipo de estadística se basa en procedimientos y técnicas que permiten identificar datos extremos de alta influencia que pudieran afectar nuestras estimaciones, así como la obtención de estimadores resistentes a la distorsión, aun cuando los datos están "contaminados" o sesgados, (Rousseeuw, 1991). En este artículo está dirigido a investigadores de la ciencia de la conducta de habla hispana. Aquí referiremos medidas de descripción alternativas útiles para el primer (medidas de localización) y segundo momento (medidas de escala) de una distribución desde el punto de vista de la estadística robusta. Respecto al segundo momento se hará énfasis en el error estándar para cubrir las potenciales necesidades de inferencia estadística del investigador (Hopkins y Glass, 1978). De manera breve describiremos la lógica detrás de cada medida alternativa y mostraremos cómo obtenerlas con el software R e IBM SPSS.

FUNDAMENTOS TEÓRICOS DEL PROCEDIMIENTO

Durante el primer momento en el análisis de una distribución, el interés es captar la acumulación central de valores en una distribución, identificando un valor típico y representativo del conjunto de datos (Press, Teukolsky, Vetterling y Flannery, 2007). La medida de localización más utilizada es la media aritmética. Sus propiedades cuantitativas, derivadas de la utilización de la magnitud conjunta de todos los valores, la hacen un estimador muy sensible a la aglomeración central de valores de la distribución (Hopkins y Glass, 1978). En condiciones de simetría y bordes ligeros de la distribución, esta medida tiene su punto de localización justo en el centro de una distribución, permitiendo estimar el valor típico y representativo en el centro de gravedad del conjunto de datos de una variable (Rosenberg y Gasko, 1983). A pesar de estas propiedades, cuando una distribución está sesgada su valor de estimación para la aglomeración central pierde precisión rápidamente. Sin embargo, un caso inquietante de distorsión en la media es cuando la distribución es simétrica, pero los bordes de la distribución son pronunciados. Se puede ilustrar esta situación utilizando la noción de distribución normal mixta o contaminada introducida por Tukey (1960).

La situación es ejemplificada por Wilcox (2017) de la siguiente manera: se puede partir de X como una variable aleatoria distribuida normalmente con la forma $\phi(x) = P(X \leq x)$. De modo adicional, contémplese una constante $K > 0$. Bajo esta situación, $\phi(x/K)$ seguiría

una distribución normal con una desviación estándar K . Por otro lado, suponga que ϵ es una constante $0 \leq \epsilon \leq 1$. Entonces la distribución normal contaminada emergerá con la siguiente forma:

$$H(x) = (1 - \epsilon) \phi(x) + \epsilon \phi(x/K)$$

La cual tendría una media de 0, pero una varianza $1 - \epsilon + \epsilon K^2$. El problema radica en que un valor diferente a 0 en ϵ causaría la presencia de una "contaminación" en forma de un incremento del grosor de los bordes de la distribución. De esta manera, tanto la estimación de localización como el desarrollo de inferencias a partir de los errores estándar de esa distribución presentarían graves problemas.

Para favorecer la comprensión, aterrizaremos la situación descrita con un ejemplo. Imagine que tiene la distribución de los puntajes totales de retención en una tarea de libre recuerdo de un conjunto de personas. Para propósitos pedagógicos, la distribución resultante está estandarizada con una media de 0 y una desviación estándar de 1. Sin embargo, una pequeña fracción de las personas que conforman nuestra población fueron abducidas por alienígenas. A pesar de que su distribución de puntajes también sigue una curva normal, en estas personas se presentan variaciones irregulares en ambos extremos de la distribución, incrementando su desviación estándar (puntajes muy altos o bajos). Si se analiza la distribución combinada sin hacer distinción de este último número de personas, los valores extremos en los bordes de la distribución reducirían de manera grave la precisión de las estimaciones basadas en la media de la distribución global. En el caso de estimaciones a partir de esa distribución, por ejemplo, se observaría que los errores estándar estarían inflados, y por tanto los intervalos de confianza producidos en la media serían anormalmente amplios, y por tanto con bajo poder (Cohen, 1988). Wilcox (2017) discute que este tipo de situaciones son difíciles de identificar, sobre todo si se considera que la distancia de Kolmogorov entre la distribución normal y la normal contaminada no es muy pronunciada.

La recurrencia de valores extremos y otras desviaciones severas en las distribuciones de datos ocurren con relativa frecuencia en la investigación del comportamiento (Field y Wilcox, 2017). ¿Qué se puede hacer al respecto? Al contar con datos derivados de la medición de alguna variable de la conducta, un primer paso es detectar la presencia de valores extremos en ambos lados de la distribución. Respecto a esta tarea, un criterio estándar es tomar como referencia el producto del rango intercuartil por 1.5 ($IQR * 1.5$). Así, cualquier valor que supere este criterio podría clasificarse como extremo y afectaría de modo potencial las estimaciones de centralidad. Al detectar outliers, el investigador puede tomar la decisión acerca de cómo proceder con base en las

características de su distribución e intereses particulares en el estudio. De manera tradicional se han reconocido dos métodos para afrontar la presencia de datos atípicos: 1) remover los outliers y usar la media aritmética de la distribución resultante como medida de localización, y 2) utilizar un estimador robusto y resistente al efecto de los outliers en el error estándar en lugar de la media aritmética (Goodall, 1983). Sin embargo, hay evidencia que demuestra las dificultades en la detección y eliminación de casos extremos (Maronna, Martin, Yohai y Salibán-Barrera, 2019). A pesar de las bondades que ofrece la detección de valores extremos, la sólida base matemática que respalda a los estimadores robustos, así como su efectividad respalda la segunda opción (Wilcox, 2017). Este manuscrito se enfocará en la segunda aproximación, en la cual se utilizan estimadores robustos alternativos a la media aritmética utilizando SPSS IBM y R como softwares estadísticos.

En seguida se mostrará cómo efectuar la detección de outliers y el cálculo de diversas medidas de localización y escala utilizando R. Previo a ilustrar estos procedimientos, vale la pena abundar en este método de análisis. R es un lenguaje de programación que permite a los investigadores de distintas disciplinas hacer análisis de datos. Al utilizar R se trabaja con un dataframe o una estructura de datos. Piense en el dataframe como una hoja de cálculo de Excel donde tiene acomodados sus datos. Cada comando que utilizaremos hará referencia al dataframe y a la variable concreta de interés en la estructura de datos. La estructura básica para analizar una variable es `dataframe$Variable`. La información básica para trabajar con dataframes la puede encontrar en <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>. Por otro lado, es importante indicar que R expande su capacidad de análisis por medio de la instalación de paquetes. Instalar un paquete permite utilizar procedimientos especiales de análisis de datos. En este artículo ilustraremos procedimientos utilizando los paquetes WRS y WRS2 (Mair y Wilcox, 2020; Wilcox, 2017; Wilcox y Schonbrodt, 2017).

DETECCIÓN DE OUTLIERS Y MEDIDAS ROBUSTAS DE LOCALIZACIÓN USANDO R

Trabajaremos la ilustración de todos los procedimientos con una pequeña distribución de datos de tiempos entre respuestas (en milisegundos) producidos por un sujeto humano en un programa concurrente independiente IV60s-IV60s (Ruiz, 2020). Imagine que a usted le interesa describir el valor típico de tiempo entre respuesta en segundos producido por un participante. La estimación precisa de dicho valor, y en general la distribución de

datos, dará una idea del efecto del programa de reforzamiento en la forma de responder del sujeto. Comencemos con el análisis. Los datos están organizados en un dataframe denominado datosIRT1. La única variable en este dataframe se titula IRT. Previo a analizar datos, es conveniente que conozca la estructura de su dataframe. Un comando muy útil para visualizar esta estructura es head(dataframe). En seguida ilustraremos su uso.

```
> head(datosIRT1)
  IRT
1 2810.79
2 2799.05
3 2810.84
4 1595.15
5 2526.77
6 3416.05
```

Como se puede apreciar, la estructura de datos sólo tiene una columna donde se almacenan los valores individuales de tiempos entre respuestas producidos en la sesión. Una vez que se conoce la estructura de la base de datos se procederá a obtener un histograma de los datos para valorar la forma de la distribución con R y detectar de manera visual la presencia de datos extremos:

```
> hist(datosIRT1$IRT, xlab = "Tiempos entre
entre respuesta ms", main = "Histograma")
```

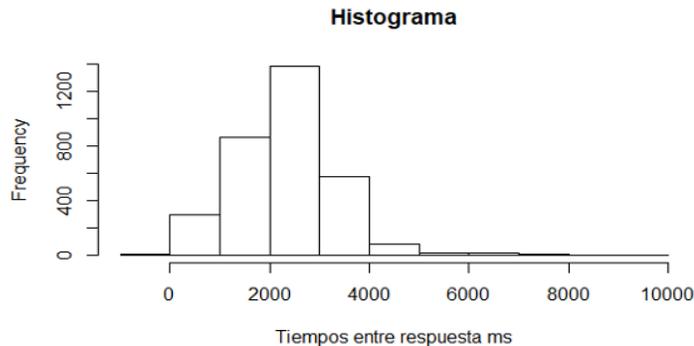


Figura 1.

Histograma de los datos de los tiempos entre respuestas proporcionado por R.

Al inspeccionar el histograma (figura 1) se detecta con rapidez la presencia de valores extremos en el lado derecho de la distribución. La presencia de dichos valores y el sesgo positivo resultante indica una posible distorsión de las estimaciones de centralidad en este conjunto particular de datos. Se procederá entonces con el proceso de detección de valores extremos (IQR * 1.5) utilizando R:

```
> datosextrmos <- boxplot.stats(datosIRT1$IRT)$out
> datosextrmos
```

```
[1] 4832.53 -222.57 4886.85 4988.27 -206.41
-280.64 4798.79 -148.72 -185.94 -79.26
[11] -723.45 4 858.52 4856.96 5528.52 -382.74
4 927.83 4684.69 6225.00 6427.00 5663.00
[21] 5694.00 5476.00 4930.00 6271.00 4882.00
6381.00 4883.00 7066.00 4883.00 7051.00
[31] 9766.00 5584.00 5522.00 7082.00 7005.00
5538.00 6942.00 6412.00 7004.00 6349.00
[41] 5538.00 4852.00 7644.00 5506.00 5054.00
6443.00 7020.00 5 710.00 4898.00 5632.00
[51] 6927.00 6224.00 5460.00 7753.00
```

El comando utiliza la función básica de estadísticas asociadas al cálculo y representación de diagramas de caja para poder computar los datos que cumplen con el criterio indicado. Los valores se almacenan en una nueva variable. El resultado del procedimiento anterior permitió conocer los valores atípicos. Dichos valores se muestran en color azul. La presencia de estos valores tiene como consecuencia que la estimación de la media aritmética resulte imprecisa para reflejar los valores típicos de la aglomeración central de los datos en la distribución de tiempos entre respuesta. Debido a esto, una opción sensible es comenzar a explorar las alternativas robustas disponibles.

Una primera opción es utilizar una media recortada (Huber, 1981; Wilcox, 2017). La media recortada consiste en estimar la media a partir de una distribución que ha sido recortada de ambos extremos. La lógica de la media recortada es: podemos comenzar partiendo de un conjunto de datos de una muestra $X_1 \dots X_n$, puestos en orden de magnitud ascendente $X(1) \leq X(2) \dots \leq X(n)$. En esta situación podemos definir la cantidad de recorte γ , contemplando que $0 \leq \gamma \leq 0.5$. Esto porque un corte de 0.5 sería equivalente a la mediana. Si tenemos en cuenta que $g = [n \gamma]$, redondeando el resultado anterior a un número entero, entonces la media recortada es:

$$X_t = \frac{X_{(g+1)} + X_{(n-g)}}{n - 2g}$$

El código para calcular una media recortada en R requiere de la instalación del paquete WRS2 (Mair y Wilcox, 2020). Una vez que el paquete es instalado, se puede calcular la media recortada con su error estándar correspondiente con el siguiente código:

```
library(WRS2)
> mean(datosIRT1$IRT, .2)
[1] 2308.121
> trimse(datosIRT1$IRT, .2)
[1] 16.56695
```

La primera línea de código llama al paquete que utilizaremos para los cálculos. Las siguientes líneas son los comandos para la media recortada y su error estándar. Los resultados aparecen en azul. Nótese que en ambas funciones hay dos argumentos. En todos los casos el primer argumento indica el conjunto de datos y la variable de donde se va a calcular el estadístico, mientras que el segundo argumento corresponde a la proporción deseada de γ (en el caso del ejemplo, el recorte fue de 0.2).

Existen otras alternativas robustas disponibles. Una de ellas es la estimación de la media winsorizada. Este estadístico recibe su nombre en honor de Charles P. Winsor (1895-1951). El cálculo de este estadístico tiene una lógica similar a la media recortada: Partiendo de la proporción delimitada en γ , los valores que entran en la proporción inferior y superior del recorte, en lugar de ser recortados, ahora son igualados al valor más chico y grande respectivamente de los valores no recortados a cada lado de la distribución. Asumiendo que el paquete WRS2 ha sido llamado, el cálculo del estadístico utilizando R es:

```
> winmean(datosIRT1$IRT, .2)
[1] 2303.937
> winse(datosIRT1$IRT, .2)
[1] 16.56013
```

Donde γ sigue siendo el argumento de la derecha en la función, el cual ahora define la porción a winsorizar.

Hay otra alternativa de estimación robusta perteneciente a la familia de los estimadores M, cuyo funcionamiento está basado en minimizar lo más posible funciones de pérdida emergente a partir de los datos (Wilcox, 2017). La lógica de un estimador M es como sigue. Recuérdese que la media aritmética de una muestra minimiza la función de pérdida de las distancias cuadráticas de los datos respecto de μ (ie. $\sum (x_i - \mu)^2$). El cálculo de un estimador M parte de la idea de cambiar la forma de la función de pérdida de diferencias cuadráticas respecto de μ a una función de distancia respecto de un valor x . En esta nueva función, la diferencia de cada valor de la distribución se plantea con base en un nuevo valor (x) con el objetivo de minimizar al máximo las distancias resultantes al resolver para (x). De esta manera se buscaría que la derivada resultante de dicha función de distancia produjera como resultado un valor de 0 a partir de estimación de x por medio de procedimientos iterativos. El valor final del parámetro x , desde el cual se calcula la distancia, corresponde al estimador M. Wilcox (2017) proporciona un tratamiento formalizado de esta definición. El siguiente código en R permite calcular el estimador M de una distribución y su error estándar:

```
> mest(datosIRT1$IRT)
[1] 2304.251
```

```
> mestse(datosIRT1$IRT)
[1] 16.47296
```

MEDIDAS ROBUSTAS DE ESCALA EMPLEANDO R

Hasta el momento se han cubierto medidas de localización y sus errores estándar asociados. Sin embargo, se pueden conseguir medidas de escala más robustas para describir el segundo momento de un conjunto de datos. Una de las medidas robustas más útiles en este contexto es el rango intercuartil. El rango intercuartil es definido como la diferencia entre $Q_3 - Q_1$, calculados a partir de la distribución de datos. Una propiedad útil del rango intercuartil es que su amplitud consigue capturar la distancia cuantitativa que tiene la aglomeración de valores comunes de la distribución, aun en condiciones de sesgo o valores extremos (Hopkins y Glass, 1978). El rango intercuartil se puede calcular fácilmente en R con el siguiente código:

```
> IQR(datosIRT1$IRT)
[1] 1188.045
```

Una modificación más robusta del rango intercuartil es su cálculo a partir de lo que se ha denominado ideal fourths o los cuartos ideales (Wilcox, 2017). Teniendo en cuenta la lógica del método estándar de cálculo de los cuartiles, las fórmulas para obtener los cuartiles inferior y superior a partir de los cuales se calcula el rango intercuartil modificado son:

Cuartil inferior: $q_i = (1 - h) X(j) + hX(j+1)$
 Cuartil Superior: $q_s = (1 - h) X(k) + hX(k-1)$

Donde $j = [(n/4) + (5/12)]$, redondeado al número entero siguiente, y $h = (n/4) + (5/12) - j$. Teniendo en cuenta lo anterior, el nuevo rango intercuartil sería $IQR = q_s - q_i$. Utilizando el paquete WRS en R, es posible calcular los cuartiles superior e inferior con el siguiente código:

```
> idealf(datosIRT1$IRT)
$ql
[1] 1710.923
$qu
[1] 2899.434
```

Por tanto, si se hace la resta entre el cuartil superior e inferior, el rango intercuartil basado en los cuartos ideales sería 1,188.511%.

El mejor método de comparación para demostrar la efectividad de los estadísticos robustos respecto a los estimadores estándar es su comparación a partir de un set de datos (Wilcox, 2017). ¿Cómo fue la distribución de tiempos entre respuestas de nuestros sujetos en el programa concurrente? En la tabla 1 se muestra la comparación entre los diferentes estimadores obtenidos con

cada uno de los procedimientos. El valor típico de tiempo entre respuesta es de 2 300 m, con una amplitud de variación cuantitativa de poco más de un segundo. En la tabla se incluyen la media, la mediana y el error estándar de los tiempos entre respuestas para su comparación con los estadísticos que hemos estado calculando. Es evidente cómo las medidas de localización robustas están más a la izquierda de la distribución, acercándose al clúster de datos y compensando el sesgo positivo causado por los valores extremos en la distribución de datos. De hecho los estadísticos robustos de localización son más cercanos a la mediana, y por tanto más cercanos al punto modal de la distribución.

Tabla 1.

Comparación de los estadísticos obtenidos a partir de la muestra de datos.

LOCALIZACIÓN		INFERENCIA	
Media	2 319.54	Error estándar (EE)	17.14
Mediana	2 304.28	EE (trim)	16.56695
Media recortada (trim)	2 308.121	EE (win)	16.56013
Media winsorizada (win)	2 303.937	EE (mest)	16.47296
Estimador M (mest)	2 304.251	Escala	
		IQR	1 188.045
		IQR (idealf)	1 188.511†

Abreviaciones: trim(recortada), win (winsorizada), mest (estimador m) y idealf (cuartiles ideales).

Por otro lado se puede observar la reducción en los errores estándar cuando se utilizan las alternativas robustas. Imagine que hace manipulaciones paramétricas en los programas concurrentes y desea comparar el valor de la media de las distribuciones de respuesta para cada programa concurrente. Las medidas robustas calculadas le permitirían obtener estimaciones de localización más acertadas para cada distribución de tiempo de respuestas, así como calcular intervalos de confianza más precisos o efectuar algún procedimiento de prueba de hipótesis en el contexto de la inferencia estadística (Casella y Berger, 2002; Huber, 1981).

ESTADÍSTICA ROBUSTA DE LOCALIZACIÓN Y ESCALA EMPLEANDO SPSS COMO SOFTWARE ESTADÍSTICO

Una vez revisados los procedimientos para obtener medidas robustas de localización y escala de un conjunto de datos con el software R, se procederá a revisar las posibilidades que ofrece el paquete estadístico IBM SPSS. Para tal propósito se usará el programa IBM SPSS Statistics en su versión 25 y la base de datos utilizada en el procedimiento efectuado en R con una única variable denominada IRT ("Tiempo entre respuestas").

Al iniciar el IBM SPSS se desplegarán dos ventanas de trabajo. La primera es la base de datos con su pestaña de "vista de datos" y "vista de variables". En la primera pestaña se puede observar la columna con la variable IRT, y en filas el número de casos o datos para esta variable. En la pestaña "vista de variables" se ofrecen en lista las variables activas en nuestra base de datos para ser configuradas, ordenadas y definidas (figura 2). Para más información acerca del manejo de bases de datos en SPSS, se puede consultar la documentación del software estadístico en el siguiente enlace <https://www.ibm.com/support/pages/node/618179>. La segunda ventana de trabajo corresponde a la hoja de resultados, la cual desplegará la información solicitada al software estadístico, desde los parámetros estadísticos, tablas y gráficos solicitados.

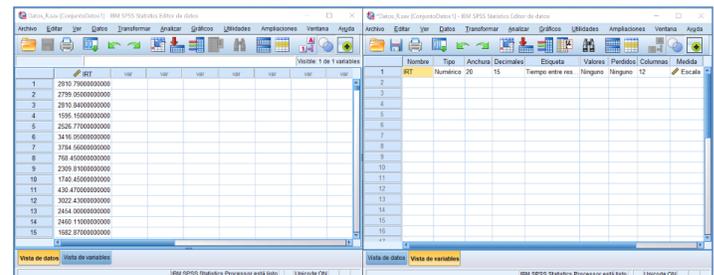


Figura 2.

Ventana de trabajo Base de Datos IBM SPSS Statistics 25.

Un primer paso es explorar la simetría de la distribución para detectar posibles sesgos que pudieran afectar las medidas de localización. Para hacer esto, el investigador puede servirse de un análisis visual de los datos por medio de un histograma. Adicional a la producción del histograma, el SPSS permite computar al mismo tiempo los estadísticos descriptivos de la distribución. Dicha solicitud puede hacer de múltiples maneras dentro del software estadístico. Para este ejercicio se mostrarán sólo dos métodos. El primer procedimiento es: Menú "Analizar", seleccionamos el submenú "Estadísti-

cos descriptivos” y después “Frecuencias”. Al ejecutar este comando se desplegará la interfaz de solicitud en la cual se deberá seleccionar la variable a analizar, en este caso IRT. Para solicitar los estadígrafos, se hace clic en “Estadísticos” y seleccionamos los que sean de nuestro interés. Para efectos prácticos se recomienda la media, mediana, desviación estándar, rango, mínimo, máximo, error estándar de la media, asimetría y cuartiles. Por otra parte, para solicitar el histograma damos clic en “Gráficos” y solicitamos la opción de histograma. La serie de selecciones se muestra en la figura 3. Una vez hecho esto, se acepta la solicitud con lo cual se desplegará el análisis en la ventana de resultados.

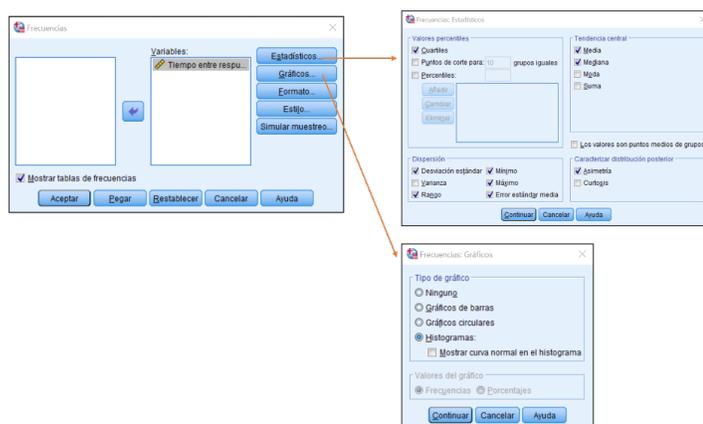


Figura 3.
Interfaz de solicitud procedimiento 1 para descriptivos e histograma.

El análisis visual exploratorio de la distribución por medio del histograma muestra datos atípicos y aislados en el lado derecho del gráfico, lo cual se puede interpretar como una distribución asimétrica positiva con la presencia de una concentración de valores extremos a la derecha de la distribución (figura 4). Al inspeccionar la tabla 2 se confirma esta situación al obtenerse un valor positivo de la asimetría (0.703), así como la ligera inconsistencia entre la media y la mediana que, en distribuciones con colas simétricas, deberían coincidir al centro de la distribución (Rosenberg y Gasko, 1983).

Tabla 2.
Estadísticos descriptivos para la variable “Tiempo” entre respuestas arrojados por el software IBM SPSS.

N	Válido	3 238
	Perdidos	0
Media		2 319.54248
Error estándar de la media		17.1368774

Mediana		2 304.285
Desv. Desviación		975.147049
Asimetría		.703
Error estándar de asimetría		.043
Rango		1 0489.4500
Mínimo		-723.45000
Máximo		9 766.00000
Percentiles	25	1 710.83000
	50	2 304.28500
	75	2 899.62250

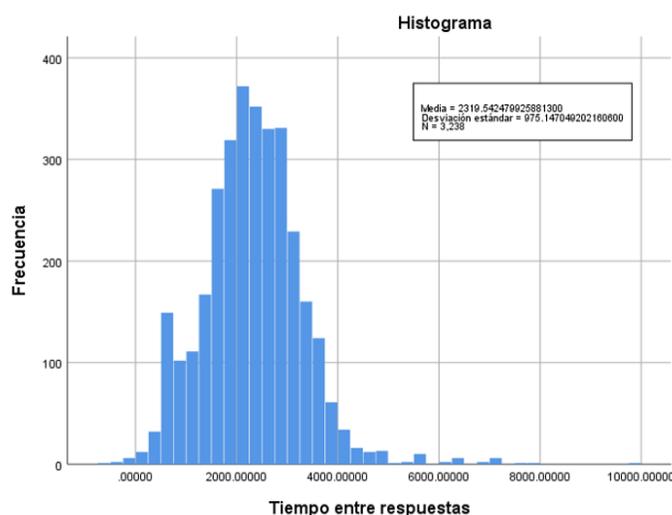


Figura 4.
Histograma de los datos de los tiempos entre respuestas proporcionado por IBM SPSS 25.

Para dar seguimiento a estos hallazgos, conviene hacer un análisis minucioso de las características de la distribución mediante un diagrama de cajas (boxplot), un gráfico de tallo y hojas, y la obtención de estimadores robustos de localización. Para hacer esta solicitud en SPSS se debe seguir la siguiente ruta: Menú “Analizar”, submenú “Estadísticos descriptivos” y se da clic en “Explorar”, lo cual desplegará una interfaz ligeramente distinta al procedimiento anterior. Tenemos que seleccionar la variable de interés a la casilla “Lista de dependientes”, y comenzamos solicitando en el apartado “Estadísticos” los descriptivos, los estimadores M y los valores atípicos. Asimismo, en el apartado “Gráficos” se seleccionarán “De tallo y hojas” e “Histograma” (figura 5).

Comenzaremos con los gráficos, analizando el gráfico de tallo y hojas solicitado (figura 6). Básicamente

procedimiento se obtuvo el rango intercuartil (tabla 2). Así, $Q_3 = 2\ 899.62250 - Q_1 = 1\ 710.8300$ arroja un rango intercuartil de $1\ 188.793$. A su vez este dato permite conocer el límite inferior y superior de nuestra distribución, así $L_i = Q_1 - 3/2 d_F$, mientras que $L_s = Q_1 + 3/2 d_F$. Todo dato fuera de estos límites correspondería a casos extremos. El segundo procedimiento proporciona por defecto el diagrama de caja, como se muestra en la figura 7. La caja o el espacio intercuartilar corresponde a 50% de la distribución alrededor de la mediana. Fuera de los bigotes se localizan los outliers con un círculo y los valores atípicos extremos con un asterisco. La diferencia entre uno y otro es que los valores atípicos extremos se alejan del cuartil inferior y superior $2\ 3/2 d_F$, y como se ha indicado, deben tratarse con mucho cuidado (Milton, 2007). Además, el diagrama de caja producido por SPSS indica el número de caso en que se presentan las anomalías para que al investigador le sea posible estimar con facilidad el grado de recorte o winsorización a efectuar en la distribución.

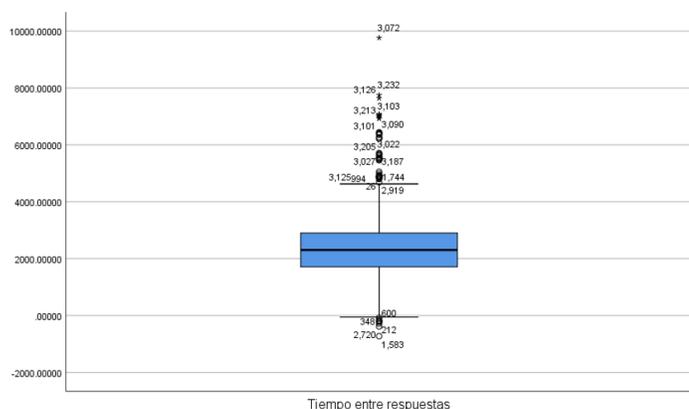


Figura 7. Gráfico de caja y bigotes para determinar localización, dispersión, asimetría, longitud de las colas y outliers de una distribución.

Una vez confirmada la presencia de valores extremos en la distribución, se puede tomar acción sobre los datos. Field (2013) indica varias acciones para reducir la influencia del sesgo en nuestros datos empleando IBM SPSS: 1) recortar la distribución; 2) winsorizar los datos para obtener la media (de manera manual), y 3) solicitar medidas robustas, como los estimadores M (calculado por SPSS). La primera opción implica recortar nuestra distribución eliminando datos extremos. En caso de recurrir al método de recorte, el investigador puede obtener una media recortada (revisada y comentada en el procedimiento empleando R). La media recortada se obtuvo con IBM SPSS mediante el segundo procedimiento revisado

(figura 5). El resultado proporciona estadísticos descriptivos, entre los que se encuentra la media recortada a 5% sin posibilidad de ajustar este valor (tabla 4). De esta manera el software calculó la media aritmética de la distribución recortada 5% a cada cola de $X_1 \dots X_n$.

Tabla 4. Identificación de valores extremos con base en una distribución estandarizada z .

OZ				
	FRECUENCIA	%	% VÁLIDO	% ACUMULADO
Outliers probables positivos	5	0.2	0.2	0.2
Outliers potenciales positivos	28	0.9	0.9	1.0
Rango normal	3 132	96.7	96.7	97.7

En lugar de recortar las colas de la distribución, una segunda opción implica winsorizar la distribución de acuerdo con un porcentaje preestablecido. Sin embargo este procedimiento es manual. Primero se transforma la distribución original a las puntuaciones estandarizadas o z -scores, dando como resultado una distribución normal con media en 0 y con desviación estándar de 1. El remplazo de valores en este procedimiento se hace teniendo en cuenta que todo aquel valor < -2.58 o $> +2.58$ de z -scores representa un outlier, mientras que los valores extremos pueden categorizarse de esta manera al encontrarse > 3.29 z -scores (Field, 2013).

La transformación a puntuaciones z y localización de extremos puede hacerse en el IBM SPSS con el siguiente procedimiento.

Transformación a puntuaciones estandarizadas. Menú "Analizar", submenú "Estadísticos Descriptivos", sin hacer cambios en los apartados de la derecha, sólo seleccionamos en la interfaz la casilla "Guardar valores estandarizados como variables". Al hacer esto se añadirá a nuestra ventana de datos una nueva columna de variable con la etiqueta Puntuación Z , en este caso ZIRT.

Categorización de extremos. Para clasificar en categorías con base en la probabilidad de casos en segmentos de la curva, iremos al menú "Transformar" y submenú "Recodificar en distintas variables", seleccionaremos las puntuaciones estandarizadas transformadas con el paso anterior, y colocaremos un nombre y etiqueta para esta nueva variable. Para determinar los valores que constituirán cada categoría damos clic en "Valores antiguos y nuevos"

y categorizamos por nivel de probabilidad de la siguiente manera: (3.2901 thru Highest=4) (2.5801 thru 3.29=3) (1.9601 thru 2.58=2) (-1.9601 thru 1.96=1) (-2.5801 thru -1.96=-2) (-3.2901 thru -2.58=-3) (Lowest thru -3.29=-4). Después se deberán nombrar las categorías con base en los valores desde la vista de variables de la siguiente manera: -4=Extremos positivos; -3=Outliers probables positivos; -2=Outliers potenciales positivos; 1=Rango normal; 2=Outliers potenciales negativos; 3=Outliers probables negativos; 4=Extremos negativos.

Identificación de extremos. Menú "Analizar", submenú "Estadísticos descriptivos", seleccionamos la opción "Frecuencias", dejamos la sección "Estadísticos" en blanco al igual que "Gráficos" y el resto de los submenús; en la interfaz seleccionamos la casilla "Mostrar tablas de frecuencias" y ejecutamos el análisis. En la hoja de resultados se observará una tabla similar a la tabla 4, en la cual se muestra la frecuencia de cada categoría y el porcentaje con relación a la distribución en su totalidad. El investigador tendrá la posibilidad de conocer el porcentaje de sus datos extremos y probables outliers que comprometen la distribución con una asimetría positiva, y por consiguiente su media aritmética.

Reemplazo de valores extremos. En IBM SPSS no hay una solicitud específica para winsorizar una distribución, por lo cual el procedimiento debe hacerse de manera manual para obtener la media winsorizada. La tabla de frecuencias obtenida en el paso anterior, en conjunto con el gráfico de tallo y hojas, así como la tabla dinámica de valores ascendentes y valores extremos, permitirán al investigador conocer el dato concreto a reemplazar. Por ejemplo, en esta distribución podrían reemplazarse los valores >4685.00 con base en el gráfico de tallo y hojas (figura 7), o los cinco valores extremos mayores y menores de la distribución con base en la tabla de valores extremos (tabla 3). Para reemplazar es necesario seleccionar la columna de la variable en cuestión en la base de datos, y dar clic en el ícono de búsqueda y reemplazo para teclear el valor específico, o ir al menú "Editar" y después "Ir al caso" para reemplazar de modo manual el valor extremo, y posteriormente volver a solicitar los estadísticos descriptivos con el primer procedimiento (figura 3), lo cual arrojará la Media de la distribución Winsorizada.

La tercera opción robusta proporcionada por el programa IBM SPSS son los estimadores M. Esta solicitud al software se hizo en el segundo procedimiento revisado antes (figura 5) al seleccionar la casilla "Estimadores M". Como se muestra en la tabla 5, el software proporciona cuatro diferentes tipos de estimadores. Los más usados son el estimador Huber y el Bponderado de Tukey. El primero está recomendado para distribuciones que se

pueden adaptar más a una distribución Gaussiana, porque pierde eficiencia en casos contrarios. En cambio, el segundo es elegido en situaciones en que hay valores extremos que pronuncian las colas de nuestra distribución. De esta manera, la función pondera con 0 los datos mientras más se alejan del estimador, con lo cual pierden su influencia (Wilcox, 2017; Goodall, 1983).

Tabla 5.

Estimadores M solicitados al explorar los datos en IBM SPSS.

	ESTIMADORES M			
	Estimador M de Huber ^a	Bponderado de Tukey ^b	Estimador M de Hampel ^c	Onda de Andrews ^d
Tiempo entre respuestas	2 306.890	2 298.375	2 298.259	2 298.018

a La constante de ponderación es 1.339.

b La constante de ponderación es 4.685.

c Las constantes de ponderación son 1.700, 3.400 y 8.500.

d La constante de ponderación es 1.340* π .

CONCLUSIÓN

La estadística robusta proporciona alternativas eficaces para producir medidas de localización y escala útiles en la descripción de una distribución con algún grado de sesgo o con la presencia de valores extremos. Goodall (1983) concluye que la eficiencia y resistencia de los estimadores robustos permite adaptarse a varias condiciones adversas. El investigador es quien puede hacer una elección adecuada teniendo en cuenta las bondades y limitaciones de cada uno de los distintos tipos de situaciones muestrales.

En el ejemplo utilizado en este artículo se obtuvieron estadísticos de localización y escala que permitieron mayor precisión en la descripción de los valores típicos de tiempos entre respuestas. También se obtuvieron errores estándar que permitían construir intervalos de confianza más pequeños para posibles comparaciones ante variaciones paramétricas en los programas concurrentes.

Como investigador debe considerar que el uso de estas alternativas puede mejorar mucho las conclusiones a las que puede llegar al momento de hacer investigación. Esto podría ser cierto en contextos aplicados donde la falta de control experimental puede generar condiciones adversas en los datos obtenidos, ¿Qué estadístico debo usar cuando mis datos se encuentren bajo las condiciones antes citadas? Wilcox (2017) reporta comparaciones entre diferentes estimadores robustos utilizando simulaciones de muestreo a partir de distintos tipos de distribuciones con bordes pronunciados (valo-

res extremos). Dos estimadores que tuvieron un mejor rendimiento y demostraron producir errores estándar relativamente pequeños en muchas situaciones fueron la media recortada a 20% y el estimador M basado en la propuesta de Huber (1981).

Estos procedimientos, en conjunto con otras posibilidades de análisis, fueron ilustrados en este artículo. Se procuró proporcionar al lector suficientes opciones de cálculo utilizando R y el software estadístico IBM SPSS. Si bien la programación en el lenguaje R ofrece más posibilidades de cálculo de medidas robustas de localización y escala, el SPSS puede servir de apoyo en la fase de exploración visual de los datos o para quienes tienen dificultades con el lenguaje de programación. El uso de una u otra alternativa queda a consideración del lector, de acuerdo con sus intereses y posibilidades.

Como conclusión, es importante destacar que la responsabilidad en la presentación y descripción de los datos recae en el investigador. La investigación científica no se debe entender como un conjunto de reglas de dedo que el psicólogo deba seguir ciegamente. En este sentido es importante que el investigador de la conducta se comprometa a la comprensión y uso razonado de los procedimientos, tomando la última decisión de acuerdo con las características de su estudio, las características de los procedimientos estadísticos disponibles y sus implicaciones en los datos. ■

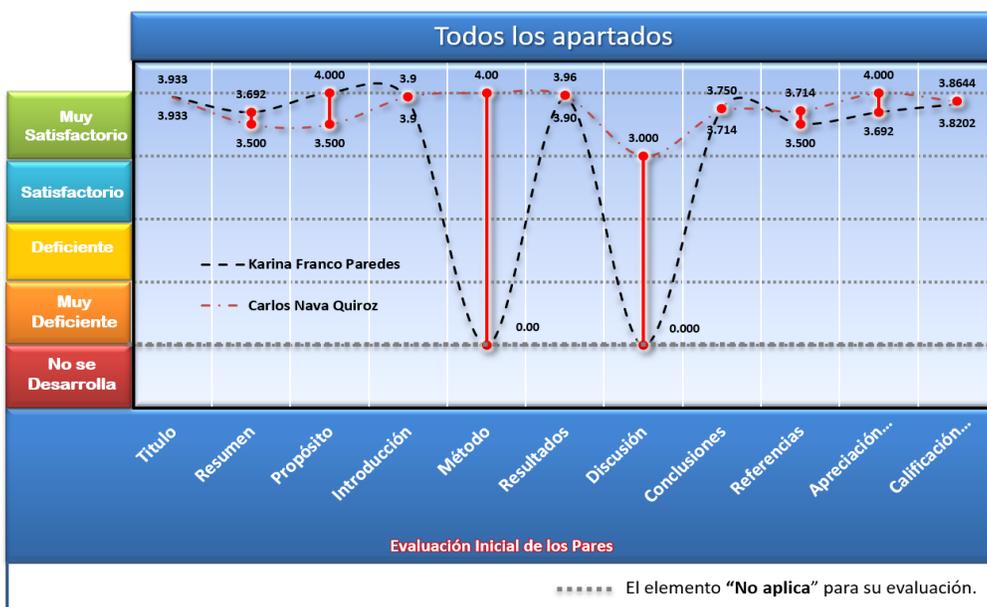
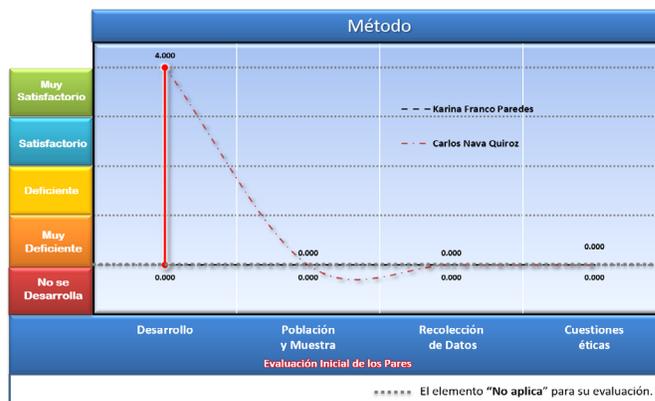
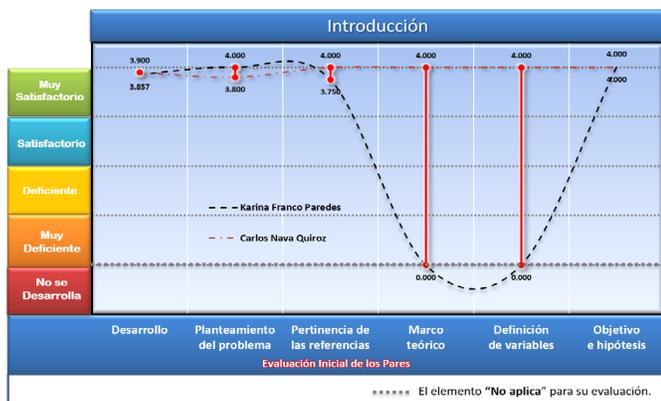
REFERENCIAS

- Casella, G., & Berger, R. (2002). *Statistical Inference*, 2a. ed. Pacific Grove.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2a. ed. Academic Press.
- Emerson, J. D., & Strenio, J. (1983). Boxplots and Batch comparison. En D. C. Hoaglin, F. Mosteller, & J. W. Tukey (eds.). *Understanding Robust and Exploratory Data Analysis*. Wiley.
- Fernández, S. F., Sánchez, J. M. C., Córdoba, A., y Largo, A. C. (2002). *Estadística descriptiva*. Esic Editorial.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.
- Field, A., & Wilcox, R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, 98(19), 38. <https://doi.org/10.1016/j.brat.2017.05.013>.
- Goodall, C. (1983). M-Estimators of location: An outline of the theory. En D. C. Hoaglin, F. Mosteller & J. W. Tukey (eds.). *Understanding Robust and Exploratory Data Analysis*. Wiley.
- Hopkins, K., & Glass, G. (1978). *Basic Statistics for the Behavioral Sciences*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Huber, P. (1981) *Robust statistics*, New York: John Wiley & Sons, Inc.
- Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavioral Research Methods*, 52(2), 464-488. doi: [10.3758/s13428-019-01246-w](https://doi.org/10.3758/s13428-019-01246-w).
- Maronna, R., Martin, D., Yohai, V., & Salibán-Barrera, M. (2019). *Robust Statistics: Theory and Methods (with R)*. John Wiley & Sons Ltd.
- Milton, J. S. (2007). *Estadística para biología y ciencias de la salud*, 3a. ed. McGraw-Hill, Inc.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*, 3a. ed. McGraw-Hill, Inc.
- Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (2007). *Numerical Recipes: The Art of Scientific Computing*, 3a. ed. Cambridge University Press.
- Rosenberg, J. L., & Gasko, M. (1983). Comparing Location Estimators: Trimmed Means, Medians, and Trimean. En D. C. Hoaglin, F. Mosteller, & J. W. Tukey (eds.). *Understanding Robust and Exploratory Data Analysis*. Wiley.
- Rousseeuw, P. J. (1991). Tutorial to robust statistics. *Journal of Chemometrics*, 5(1), 1-20. <https://doi.org/10.1002/cem.1180050103>.
- Ruiz, D. (2020). *Distribución de la conducta operante en humanos: Desarrollo de preferencia y parámetros de reforzamiento relevantes*. Tesis doctoral no publicada. Universidad Nacional Autónoma de México.
- Tukey, J. W. (1960). A survey of sampling from contaminated normal distributions. En I. Olkin, S. Ghurye, W. Hoeffding, W. Madow & H. Mann (eds.). *Contributions to Probability and Statistics*. Stanford. Stanford University Press.
- Wilcox, R. (2017). *Introduction to Robust Estimation & Hypothesis Testing*, 4a. ed. Elsevier.
- Wilcox, R., & Schonbrodt, F. (2017). *A Package of R. R. Wilcox' Robust Statistics Functions*. R package version 0.34, URL. <https://github.com/nicebread/WRS/tree/master/pkg>.

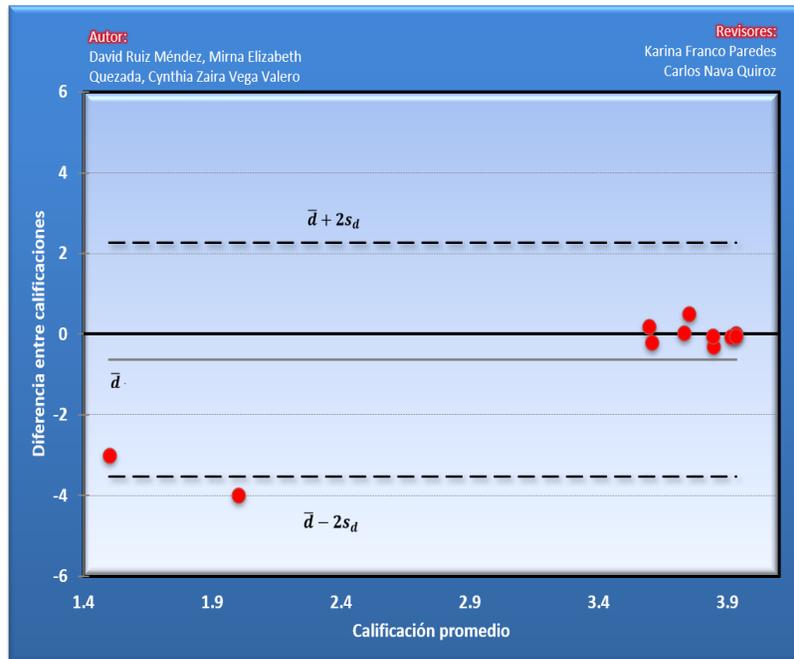


DIMENSIÓN CUANTITATIVA

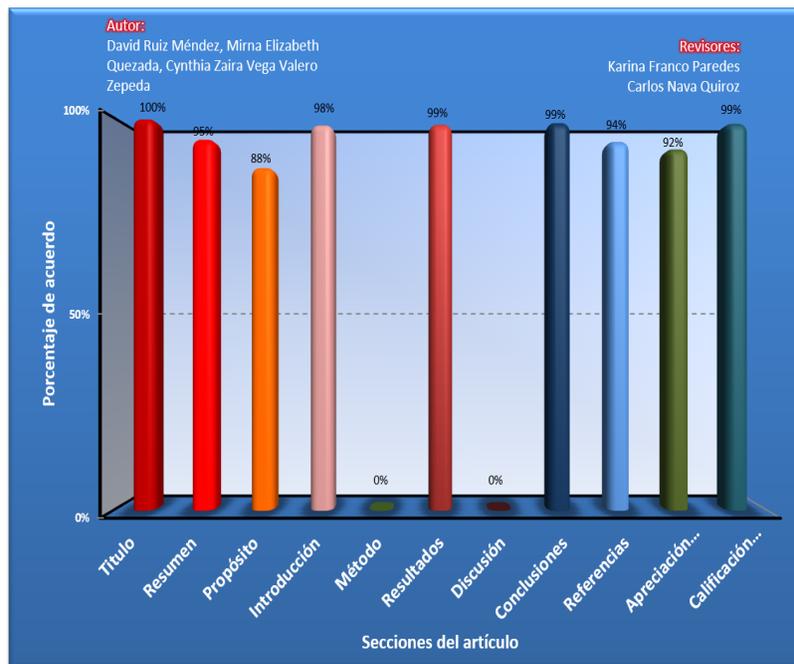
Perfil de Evaluación entre pares



Índice de Concordancia



Índice de Acuerdo



DIMENSIÓN CUALITATIVA

Revisor 1	Revisor 2
Karina Franco Paredes	Carlos Nava Quiroz
Título/Autoría	
El título es correcto	Sin comentarios
Resumen	
Ninguno	En general, el resumen, tanto en inglés como español, son satisfactorios.
Propósito del Estudio	
El propósito es claro	Es importante que el lector tenga claro a quién va dirigido el escrito, por lo cual es posible afirmar que a todos aquellos que deseen analizar datos (estudiantes e investigadores) y que posean conocimientos de estadística descriptiva e inferencial.
Introducción	
La introducción es muy puntual y fundamenta claramente el propósito de la nota técnica	El escrito es de carácter metodológico por lo que no contiene hipótesis, en cambio presenta estrategias metodológicas novedosas que son de interés, y ayudarán a los especialistas, como apoyo en el análisis de datos.
Método	
No aplica esta evaluación dado que se trata de una nota técnica	Es un artículo teórico metodológico
Resultados	
Se sugiere que los autores consideren la figura violinPlot que permite una clara visualización de la distribución de los datos e incluye un marcador para la mediana de los datos y un cuadro que indica el rango intercuartil.	Sería conveniente que presentaran un ejercicio (clínico o de investigación) que muestre las bondades de las técnicas de análisis propuestas -todo ello sin hacer alusión a justificaciones teóricas que ya fueron hechas en apartados anteriores-

Revisor 1	Revisor 2
Discusión	
Dada la naturaleza del manuscrito no se presenta esta sección	Los índices propuestos están al margen de una discusión teórico - empírica de aplicación, debido a que el escrito es de carácter metodológico- estadístico
Conclusiones	
La conclusión es adecuada, no obstante, considero que es necesario argumentar más acerca de las implicaciones para el tratamiento estadístico de los datos, iniciando el análisis de los datos con estadística robusta.	Creo que con la presentación de un ejercicio (que sea autocontenido) sería suficiente para que los lectores terminen de comprender la importancia de los planteamientos analíticos propuestos.
Referencias	
Se recomienda revisar el formato de las referencias, por ejemplo, el nombre de algunas revistas no está en cursiva, el título de la investigación está entre comillas, etc	Es probable que pudieran incluirse citas donde se apliquen las técnicas propuestas, como ejemplos de aplicación.

