GUÍA DE PRÁCTICA

MEDIDAS DE POSICIÓN EN VARIABLE CONTINUA Y DISCRETA – DATOS DESAGRUPADOS CON MICROSOFT EXCEL

Diana María Robayo-Botiva
Universidad Cooperativa de Colombia
Sede Villavicencio

Apropiación social del conocimiento

Generación de contenidos impresos

<http://repository.ucc.edu.co/handle/20.500.12494/7375>

N.º 19, septiembre de 2020

doi: <https://doi.org/10.16925/gcgp.26>

NOTA LEGAL

El presente documento de trabajo ha sido incluido dentro de nuestro repositorio institucional como Apropiación social de conocimiento por solicitud del autor, con fines informativos, educativos o académicos. Asimismo, los argumentos, datos y análisis incluidos en el texto son responsabilidad absoluta del autor y no representan la opinión del Fondo Editorial o de la Universidad.

DISCLAIMER

This coursework paper has been uploaded to our institutional repository as Social Appropriation of Knowledge due to the request of the author. This document should be used for informational, educational or academic purposes only. Arguments, data and analysis included in this document represent authors' opinion not the Press or the University.

EDICIONES

Universidad Cooperativa
de Colombia

GUÍA DE PRÁCTICA

**MEDIDAS DE POSICIÓN
EN VARIABLE CONTINUA
Y DISCRETA – DATOS
DESAGRUPADOS CON
MICROSOFT EXCEL**

Diana María Robayo-Botiva

ACERCA DEL AUTOR

Diana María Robayo-Botiva, magíster en comercio electrónico del Instituto Tecnológico y de Estudios Superiores de Monterrey y magíster en gestión de las tecnologías de la información de la Universidad Cooperativa de Colombia. Profesora, Facultad de Ciencias Económicas, Administrativas y Contables, Universidad Cooperativa de Colombia, sede Villavicencio, Colombia.

Correo electrónico:

diana.robayo@campusucc.edu.co

Google Scholar: <https://scholar.google.com/citations?user=XL1poXcAAAAJ&hl=es>

CÓMO CITAR ESTE DOCUMENTO

Robayo-Botiva, D. M. (2020). *Medidas de posición en variable continua y discreta – datos desagrupados con Microsoft Excel* (Generación de contenidos impresos N.º 19). Bogotá: Ediciones Universidad Cooperativa de Colombia. doi: <https://doi.org/10.16925/gcgp.26>

Este documento puede ser consultado, descargado o reproducido desde nuestro repositorio institucional (<http://repository.ucc.edu.co/handle/20.500.12494/7369>) para uso de sus contenidos, bajo la licencia de Creative Commons Reconocimiento-No Comercial-Sin Obra Derivada 4.0 Internacional. <http://creativecommons.org/licenses/by-nc-nd/4.0/>

TABLA DE CONTENIDO

INTRODUCCIÓN	6
PROPÓSITO	6
MEDIDAS DE POSICIÓN: CUARTILES, DECILES Y PERCENTILES	6
Cuartiles	6
Ejemplo datos cuantitativos continuos	6
Diagramas de Caja	7
Deciles	9
Percentiles	10
OBJETIVOS	13
MATERIALES	13
Descripción de actividades y procedimientos de la práctica en Microsoft Excel	13
Diagrama de cajas	16
TRABAJO PARA LOS ESTUDIANTES	25
REFERENCIAS	26

19 MEDIDAS DE POSICIÓN EN VARIABLE CONTINUA Y DISCRETA – DATOS DESAGRUPADOS CON MICROSOFT EXCEL

Diana María Robayo-Botiva

Resumen

En esta guía se desarrollarán ejercicios de contexto de medidas de posición para datos desagrupados de variable cuantitativa discreta y continua. Se validará el resultado del desarrollo paso a paso de cada uno de los ejercicios resueltos con el procedimiento realizado a través de la herramienta ofimática Microsoft Excel. De esta manera, el estudiante tendrá una aplicación de apoyo para la verificación de los ejercicios propuestos en la última sección de la guía. El uso de Microsoft Excel en el curso de estadística descriptiva se debe a que esta se encuentra instalada en la mayoría de los laboratorios universitarios y en la mayoría de los equipos del hogar, facilitando que los estudiantes puedan acceder a ella. Adicionalmente, muchos autores consideran que Microsoft Excel es una excelente herramienta para la enseñanza de la estadística, ya que “una ventaja importante de la hoja de cálculo Excel es que se ha convertido en un software estándar en los entornos de enseñanza, profesionales y familiares. Excel tiene una interfaz amigable y es fácil usar” (Cao y Naya, 2013 citado por Martín, 2016).

Palabras clave: cuartiles, deciles, fórmulas de Excel, percentiles.

INTRODUCCIÓN

Dentro del contenido del curso de estadística descriptiva se encuentra la temática de medidas de posición. Por lo tanto, para que el estudiante tenga una mejor comprensión, se usará Microsoft Excel. En consecuencia, en la práctica se propone, inicialmente, el desarrollo paso a paso de los ejercicios hasta encontrar la solución y luego se realizará, igualmente, el paso a paso mediante las funciones de Microsoft Excel, de tal manera que el estudiante pueda corroborar la solución de los ejercicios realizados en la sección teórica.

PROPÓSITO

Practicar las temáticas de medidas de posición: cuartiles, deciles, percentiles para datos desagrupados en variable cuantitativa continua y discreta del curso de estadística descriptiva, mediante ejercicios de contexto, utilizando la herramienta Microsoft Excel.

MEDIDAS DE POSICIÓN: CUARTILES, DECILES Y PERCENTILES

Cuando la distribución contiene un número alto de intervalos o marcas de clase y se requiere obtener un promedio de una parte de ella, se puede dividir la distribución en cuatro, en diez o en cien partes. En el primer caso se habla de cuartiles, en el segundo de deciles y, en el último, de centiles o percentiles (Martínez, Estadística y Muestreo, 2012).

Los cuartiles, deciles y percentiles se emplean con más frecuencia en variables continuas, siempre que el número de intervalos sea grande y se desee obtener un promedio que corresponda a una determinada parte de la distribución (Martínez, 2016).

Adicionalmente, las medidas de posición, también llamadas cuantiles, son aquellas que

permiten calcular valores en la distribución de los datos y que la dividen en partes iguales, de tal forma que los intervalos generados por los cuantiles contengan el mismo número de datos. Los cuantiles más usados son los cuartiles, deciles y percentiles. Cuando se tienen datos agrupados en intervalos, estas medidas se consideran en cierta forma como una extensión de la mediana (Hernández, 2016).

CUARTILES

Los cuartiles (Q_k) son valores que fraccionan la distribución de los datos en cuatro partes iguales (Ruíz Muñoz, 2005 citado en Hernández, 2016). El cuartil inferior (Q_1) es aquel valor de la variable que supera el 25% de las observaciones y a la vez, es superado por el restante 75%.

El segundo cuartil (Q_2) es aquel valor de la variable que supera el 50% y a la vez es superado por el otro 50% de las observaciones (igual a la mediana). El tercer cuartil (Q_3) es aquel valor de la variable que supera al 75% y es superado por el restante 25% de las observaciones (Martínez, 2016).

Para entender el cálculo de los cuartiles, se realizará el procedimiento a través de un ejemplo de datos desagrupados tanto para variable cuantitativa continua como cuantitativa discreta, como se relaciona a continuación.

Ejemplo datos cuantitativos continuos

La talla normal de recién nacidos a término en un hospital en una mañana fue: 46, 48, 46, 50, 53, 51, 47, 47, 51, 52, 50, 48, 49, 50, 51. Sobre esto, determinar el cuartil 2.

Para el desarrollo del ejercicio anterior realizamos el siguiente procedimiento:

Cuartil 2 (Q_2):

Se ordenan los datos de menor a mayor

46 46 47 47 48 48 49 50 50 50 51 51 51 52 53

Se calcula la posición con la ecuación $\frac{k(n+1)}{4}$ donde k es el número de cuartil ($k=1,2,3$) y n es el número total de datos.

$$\frac{2(15+1)}{4} = 8$$

En el ejemplo, $k=2$ indica el cuartil 2, $n=15$ debido a que son 15 datos en total. El cuatro nos indica que la distribución se divide en cuatro partes iguales. Finalmente, el 8 nos indica que el Q_2 corresponde al valor ubicado en la octava posición.

Si a la organización de los datos agregamos la posición:

46 46 47 47 48 48 49 50 50 51 51 51 52 53
1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a

Entonces, se concluye que el valor de Q_2 es 50 (resaltado en verde) debido a que se encuentra en la octava posición. Por lo tanto, su interpretación significa que el 50% de los recién nacidos presentaron una talla máxima de 50 cm, también se puede interpretar como el 50% de los recién nacidos tienen tallas entre 46 cm y 50 cm.

Ejemplo datos cuantitativos discretos

A continuación, se relaciona el número de fallas de 15 estudiantes de un curso de matemáticas de secundaria en 20 días de clase:

	No. de fallas
Estudiante 1	5
Estudiante 2	3
Estudiante 3	6
Estudiante 4	6
Estudiante 5	2
Estudiante 6	1
Estudiante 7	1
Estudiante 8	7
Estudiante 9	5
Estudiante 10	5
Estudiante 11	8
Estudiante 12	2
Estudiante 13	3
Estudiante 14	4
Estudiante 15	1

Sobre esto, determinar el cuartil 3.

Cuartil 3 (Q_3):

Se ordenan los datos de menor a mayor el número de fallas por estudiante.

1 1 1 2 2 3 3 4 5 5 5 6 6 7 8

Se calcula la posición con la ecuación $\frac{k(n+1)}{4}$ donde k es el número de cuartil ($k=1,2,3$) y n es el número total de datos.

$$\frac{3(15+1)}{4} = 12$$

$k=3$ indica que es el cuartil 3, $n=15$ debido a que son 15 datos en total. El cuatro nos indica que la distribución se divide en cuatro partes iguales. Finalmente, el 12 nos indica que el Q_3 corresponde al valor ubicado en la decimosegunda posición.

Si a la organización de los datos agregamos la posición:

1 1 1 2 2 3 3 4 5 5 5 6 6 7 8
1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a

Entonces, se puede concluir que el valor de Q_3 es 6 (resaltado en verde) debido a que se encuentra en la decimosegunda posición. Por lo tanto, su interpretación significa que el 75% de los estudiantes del curso de matemáticas presentaron un número de fallas máxima de 6, también se puede interpretar como el 75% de los estudiantes del curso de matemáticas tienen entre 1 y 6 fallas.

Diagrama de cajas

Es una representación gráfica basada en cuartiles que ayuda a presentar un conjunto de datos. Para construir un diagrama de cajas solo se necesitan cinco estadísticos: Valor mínimo, Q_1 (cuartil 1), Q_2 (cuartil 2 – mediana), Q_3 (cuartil 3) y valor máximo (Lind, Marchal y Wathen, 2015).

Para aprender a realizar el diagrama de cajas, se utilizará el ejemplo propuesto para variable cuantitativa continua:

La talla normal de recién nacidos a término en un hospital en una mañana fue: 46, 48, 46, 50, 53, 51, 47, 47, 51, 52, 50, 48, 49, 50, 51.

Para este ejercicio, se ha determinado el $Q_2 = 50$. Sin embargo, se debe determinar el cuartil 1 y cuartil 3. A continuación, se realizará el procedimiento para los cuartiles faltantes:

Cuartil 1 (Q_1)

Se calcula la posición del cuartil 1:

$$\frac{1(15+1)}{4} = 4$$

El resultado es 4 nos indica que el Q_1 corresponde al valor ubicado en la cuarta posición.

46	46	47	47	48	48	49	50	50	50	51	51	51	52	53
1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	13a	14a	15a

El valor de Q_1 es 47 (resaltado en verde). Su interpretación significa que el 25% de los recién nacidos presentaron una talla máxima de 47 cm, también se puede interpretar como que el 25% de los recién nacidos tienen tallas entre 46 cm y 47 cm.

Cuartil 3 (Q_3)

Se calcula la posición del cuartil 3:

$$\frac{3(15+1)}{4} = 12$$

El resultado es 12 y nos indica que el Q_3 corresponde al valor ubicado en la duodécima posición.

46	46	47	47	48	48	49	50	50	50	51	51	51	52	53
1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	13a	14a	15a

El valor de Q_3 es 51 (resaltado en verde). Su interpretación significa que el 75% de los recién nacidos presentaron una talla máxima de 51 cm, también se puede interpretar como que el 75% de los recién nacidos tienen tallas entre 46 cm y 51 cm.

Debido a que se tienen los 3 cuartiles, se graficará el diagrama a partir de los datos necesarios para este:

Valor mínimo = 46 (número menor de los datos de la distribución)

$Q_1 = 47$

Mediana o $Q_2 = 50$

$Q_3 = 51$

Valor máximo = 53 (número mayor de la distribución)

Luego, se debe crear una escala adecuada a lo largo del eje horizontal (o vertical), posteriormente, se debe dibujar una caja que inicie en Q_1 (47 cm) y termine en Q_3 (51 cm). Dentro de la caja trace una línea vertical para presentar la mediana o Q_2 (50 cm). Por último, prolongue líneas horizontales (o verticales) a partir de la caja al valor mínimo (46 cm) y al valor máximo (53 cm).

Estas líneas que salen horizontales que salen de la caja, a veces reciben el nombre de bigotes, en virtud de su parecido a los bigotes de un gato (Lind, Marchal y Wathen, 2015).

FIGURA 1. Diagrama de cajas ejemplo de variable cuantitativa continua.

Fuente: elaboración propia.

El diagrama de caja también muestra el rango intercuartil de las tallas de los recién nacidos entre Q_1 y Q_3 . Este es 4 cm e indica que el 50% de las tallas se encuentran entre 47 cm y 51 cm.

DECILES

Los deciles (D_k) son valores que fraccionan la distribución de los datos en diez partes iguales (Ruíz Muñoz, 2005 citado en Hernández, 2016). En la distribución se presentan nueve deciles: el D_1 acumula el 10% del conjunto de datos, el D_2 deja el 20%, y así sucesivamente hasta el D_9 , que acumula el 90% de los datos (Hernández, 2016).

Para entender el cálculo de los deciles, nuevamente, se realizará el procedimiento a través de los ejemplos propuestos en cuartiles:

Ejemplo datos cuantitativos continuos

La talla normal de recién nacidos a término en un hospital en una mañana fue: 46, 48, 46, 50, 53, 51, 47, 47, 51, 52, 50, 48, 49, 50, 51. Sobre estos, determinar el Decil 4.

El desarrollo del ejercicio en deciles se plantea de manera similar al de los cuartiles como se refleja a continuación:

Decil (D_4)

Se ordenan los datos de menor a mayor

46 46 47 47 48 48 49 50 50 50 51 51 51 52 53

Se calcula la posición con la ecuación $\frac{k(n+1)}{10}$ donde k es el número de cuartil ($k=1, 2, 3, 4, 5, 6, 7, 8, 9$) y n es el número total de datos.

$$\frac{4(15+1)}{10} = 6.4$$

$k=4$ indica que es el decil 4, $n=15$ debido a que son 15 datos en total. El diez nos indica que la distribución se divide en diez partes iguales. Finalmente el 6.4 nos indica que el D_4 corresponde al valor ubicado entre la sexta y séptima posición (entre 48 y 49).

Si a la organización de los datos agregamos la posición:

46 46 47 47 48 48 49 50 50 50 51 51 51 52 53
1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a

Se puede evidenciar que el decil 4 está entre dos números (resaltados en verde). Ya que el resultado de la posición no es un número entero, se desarrolla el siguiente procedimiento adicional:

Se realiza la diferencia entre el dato de la séptima y sexta posición: $49-48 = 1$ y este valor se multiplica por el excedente de 6.4, el cual es 0.4, así $1*(0.4)=0.4$ y, finalmente, este valor se agrega a 48 que se encuentra en la sexta posición, es decir, $48+0.4=48.4$.

Se concluye, entonces, que el valor de D_4 es 48.4, ya que está entre la sexta y séptima posición. Por lo tanto, su interpretación significa que el 40% de los recién nacidos presentaron una talla máxima de 48.4 cm. De igual forma, se puede interpretar como que el 40% de los recién nacidos tienen tallas entre 46 cm y 48.4 cm.

Ejemplo datos cuantitativos discretos

A continuación, se relaciona el número de fallas de 15 estudiantes de un curso de matemáticas de secundaria en 20 días de clase:

	No. de fallas
Estudiante 1	5
Estudiante 2	3
Estudiante 3	6
Estudiante 4	6
Estudiante 5	2
Estudiante 6	1
Estudiante 7	1
Estudiante 8	7
Estudiante 9	5
Estudiante 10	5
Estudiante 11	8
Estudiante 12	2
Estudiante 13	3
Estudiante 14	4
Estudiante 15	1

Sobre esto, determine el decil 7.

Decil 7 (D_7)

Se ordenan los datos de menor a mayor el número de fallas por estudiante.

1 1 1 2 2 3 3 4 5 5 5 6 6 7 8

Se calcula la posición con la ecuación $\frac{k(n+1)}{10}$ donde k es el número de decil ($k=1, 2, 3, 4, 5, 6, 7, 8, 9$) y n es el número total de datos.

$$\frac{7(15+1)}{10} = 11.2$$

$k=7$ indica que es el decil 7, $n=15$ debido a que son 15 datos en total. El diez nos indica que la distribución se divide en diez partes iguales. Finalmente, el 11.2 nos indica que el D_7 corresponde al valor ubicado entre la decimoprimer y decimosegunda posición (entre 5 y 6).

Si a la organización de los datos agregamos la posición:

1	1	1	2	2	3	3	4	5	5	5	6	6	7	8
1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	13a	14a	15a

Se puede evidenciar que el D_7 está entre dos números (resaltados en verde). Ya que el resultado de la posición no es un número entero, se desarrolla el siguiente procedimiento adicional:

Realizamos la diferencia entre el dato de la decimosegunda y decimoprimer posición: $6-5=1$ y este valor se multiplica por el excedente de 11.2, el cual es 0.2, así $1*(0.2)=0.2$ y, finalmente, este valor se agrega a 5 que se encuentra en la decimoprimer posición, es decir, $5+0.2=5.2$.

Se concluye, entonces, que el valor de D_7 es 5.2, ya que esta entre la decimoprimer y decimosegunda posición. Por lo tanto, su interpretación significa que el 70% de los estudiantes de matemáticas presentaron un numero de fallas máxima de 5.2, pero como es un dato discreto, también se podría indicar que los estudiantes presentaron un numero de fallas máxima de aproximadamente 6. De igual forma, se puede interpretar como que el 70% de los estudiantes de matemáticas tienen entre 1 y 5.2 fallas. No obstante, como son datos discretos se podría interpretar como que el 70% de los estudiantes de matemáticas tienen entre 1 y aproximadamente 6 fallas.

PERCENTILES

Los percentiles (P_k) son valores que fraccionan la distribución de los datos en cien partes iguales con igual número de observaciones (Ruíz Muñoz, 2005 citado en Hernández, 2016). Es decir, se tendrán 99 valores de las variables que separa a la frecuencia total de la distribución divididas en 100 partes iguales (Martínez, 2016).

En la distribución se presentan 99 percentiles: el primer percentil P_1 acumula el 1% del conjunto de datos, el percentil P_2 deja el 2%, y de forma similar los demás percentiles hasta llegar al percentil P_{99} que acumula el 99% de los datos (Hernández, 2016).

Para entender el cálculo de los percentiles, de nuevo se realizará el procedimiento a través de los ejemplos propuestos para cuartiles y deciles.

Ejemplo datos cuantitativos continuos

La talla normal de recién nacidos a término en un hospital en una mañana fue: 46, 48, 46, 50, 53, 51, 47, 47, 51, 52, 50, 48, 49, 50, 51. Sobre esto, determinar el percentil 90.

El desarrollo del ejercicio en percentiles se plantea de manera similar al de deciles, como se relaciona a continuación:

Percentil 85 (P_{85})

Se ordenan los datos de menor a mayor

46 46 47 47 48 48 49 50 50 50 51 51 51 52 53

Se calcula la posición con la ecuación $\frac{k(n+1)}{100}$ donde k es el número de cuartil ($k=1, 2, 3, 4, 5, 6, 7, \dots, 13, 14, \dots, 98, 99$) y n es el número total de datos.

$$\frac{85(15+1)}{100} = 13.6$$

$k=85$ indica que es el percentil 85, $n=15$ debido a que son 15 datos en total. El cien nos indica que la distribución se divide en cien partes iguales. Finalmente, el 13.6 nos indica que el P_{85} corresponde al valor ubicado entre la decimotercera y decimocuarta posición (entre 51 y 52).

Si a la organización de los datos agregamos la posición:

46 46 47 47 48 48 49 50 50 50 51 51 51 52 53
1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a

Se puede evidenciar que el percentil 85 está entre dos números (resaltados en verde). Ya que el resultado de la posición no es un número entero, se desarrolla el siguiente procedimiento adicional.

Realizamos la diferencia entre el dato de la decimocuarta y decimotercera posición: $52-51=1$. Este valor se multiplica por el excedente de 13.6, que es 0.6, así $1*(0.6)=0.6$ y, finalmente, este valor se agrega a 51, que se encuentra en la decimotercera, es decir, $51+0.6=51.6$.

Se concluye, entonces, que el valor de P_{85} es 51.6, ya que está entre la decimotercera y decimocuarta posición. Por lo tanto, su interpretación significa que el 85% de los recién nacidos presentaron una talla máxima de 51.6 cm. De igual forma, se puede interpretar como que el 85% de los recién nacidos tienen tallas entre 46 cm y 51.6 cm.

Ejemplo datos cuantitativos discretos

A continuación, se relacionan el número de fallas de 15 estudiantes de un curso de matemáticas de secundaria en 20 días de clase:

	No. de fallas
Estudiante 1	5
Estudiante 2	3
Estudiante 3	6
Estudiante 4	6
Estudiante 5	2
Estudiante 6	1
Estudiante 7	1
Estudiante 8	7
Estudiante 9	5
Estudiante 10	5
Estudiante 11	8
Estudiante 12	2
Estudiante 13	3
Estudiante 14	4
Estudiante 15	1

Sobre esto, determine el percentil 90.

Percentil 90 (P_{90})

Se ordenan los datos de menor a mayor el número de fallas por estudiante.

1 1 1 2 2 3 3 4 5 5 5 6 6 7 8

Se calcula la posición con la ecuación $\frac{k(n+1)}{100}$ donde k es el número de percentil ($k=1, 2, 3, 4, 5, 6, 7, \dots, 13, 14, \dots, 98, 99$) y n es el número total de datos.

$$\frac{90(15+1)}{100} = 14.4$$

$k=90$ indica que es el percentil 90, $n=15$ debido a que son 15 datos en total. El cien nos indica que la distribución se divide en cien partes iguales. Finalmente, el 14.4 nos indica que el P_{90} corresponde al valor ubicado entre la decimocuarta y decimoquinta posición (entre 7 y 8).

Si a la organización de los datos agregamos la posición:

1 1 1 2 2 3 3 4 5 5 5 6 6 7 8
1a 2a 3a 4a 5a 6a 7a 8a 9a 10a 11a 12a 13a 14a 15a

Se puede evidenciar que el percentil 90 está entre dos números (resaltados en verde). Ya que el resultado de la posición no es un número entero, se desarrolla el siguiente procedimiento adicional:

Realizamos la diferencia entre el dato de la decimoquinta y decimocuarta posición: $8-7=1$. Este valor se multiplica por el excedente de 14.4, que es 0.4, así $1*(0.4)=0.4$ y, finalmente, este valor se agrega a 7, que se encuentra en la decimoprimera posición, es decir, $7+0.4=7.4$.

Se concluye, entonces, que el valor de P_{90} es 7.4, ya que está entre la decimocuarta y decimoquinta posición. Por lo tanto, su interpretación significa que el 90% de los estudiantes de matemáticas presentaron un número de fallas máxima de 7.4, pero como es un dato discreto, también se podría indicar que los estudiantes presentaron

un número de fallas máxima de aproximadamente 8. De igual forma, se puede interpretar como que el 70% de los estudiantes de matemáticas tienen entre 1 y 7.4 fallas. No obstante, como son datos discretos, se podría interpretar como que el 70% de los estudiantes de matemáticas tienen entre 1 y aproximadamente 8 fallas.

Ejemplo adicional datos desagrupados variable cuantitativa continua

Para finalizar, se realizará un ejercicio con el fin de mejorar la comprensión de la temática.

A continuación, se registran las edades de mujeres del municipio de Villavicencio que han padecido Covid-19 por una fuente relacionada, que su estado ha sido leve y que se han recuperado (Instituto Nacional de Salud, 2020).

71 37 32 27 49 50 40 23 22 40 27 22 37
33 27 20 11 31 38 19 35 21 31 19 19 32

Halle el Q_2 (cuartil 2), D_8 (decil 8), P_{95} (percentil 95).

Antes de iniciar, se procederá a organizar las edades de menor a mayor:

11	19	19	19	20	21	22	22	23	27	27	27	31
1a	2a	3a	4a	5a	6a	7a	8a	9a	10a	11a	12a	13a
31	32	32	33	35	37	37	38	40	40	49	50	71
14a	15a	16a	17a	18a	19a	20a	21a	22a	23a	24a	25a	26a

Cuartil 2 (Q_2)

Utilizamos la ecuación $\frac{k(n+1)}{4}$ para determinar la ubicación del cuartil 2, en este caso sería:

$$\frac{2(26+1)}{4} = 13.5$$

Esto quiere decir que el Cuartil 2 está entre la decimotercera y decimocuarta posición, o sea 31, ya que estas posiciones tienen el mismo número. Se realizará el procedimiento descrito en ejercicios anteriores para encontrar el valor que, en este caso, sería la diferencia entre la decimotercera y la decimocuarta posición $(31-31)=0$. Este valor se multiplicaría por el excedente de 13.5, que es 0.5. Sin embargo, como es cero, entonces

sería $31+0=31$. Por lo que $Q_2 = 31$. Su interpretación indicaría que el 50% de las mujeres que tuvieron Covid-19 de tipo relacionado, en estado leve y que se han recuperado presentaron una edad máxima de 31 años.

Decil 8 (D_8)

Utilizamos la ecuación $\frac{k(n+1)}{10}$ para determinar la ubicación del decil 8, en este caso sería:

$$\frac{8(26+1)}{10} = 21.6$$

Esto quiere decir que el D_8 está entre la vigesimoprimera y vigesimosegunda posición, es decir entre las edades de 38 y 40. Se realizará, entonces, el procedimiento descrito en ejercicios anteriores para encontrar el valor que, en este caso, sería la diferencia entre la vigesimosegunda y la vigesimoprimera posición: $(40-38)=2$. Este valor se multiplica por el excedente de 21.6, el cual es 0.6, así $2*(0.6)=1.2$ y, finalmente, este valor se agrega a 38, que se encuentra en la vigesimoprimera posición, es decir, $38+1.2=39.2$. Por lo tanto, $D_8 = 39.2$. Su interpretación indicaría que el 80% de las mujeres que tuvieron Covid-19 de tipo relacionado, en estado leve y que se han recuperado presentaron una edad máxima de 39.2 años.

Percentil 95 (P_{95})

Utilizamos la ecuación $\frac{k(n+1)}{100}$ para determinar la ubicación del percentil 95, en este caso sería:

$$\frac{95(26+1)}{100} = 25.65$$

Esto quiere decir que el P_{95} está entre la vigesimoquinta y vigesimosexta posición, es decir, entre las edades de 50 y 71. Se realizará, entonces, el procedimiento descrito en ejercicios anteriores para encontrar el valor que, en este caso, sería la diferencia entre la vigesimosexta y la vigesimoquinta posición: $(71-50)=21$. Este valor se multiplica por el excedente de 25.65, que es 0.65, así $21*(0.65)=13.65$ y, finalmente, este valor se agrega a 50 que se encuentra en la vigesimoquinta posición, es decir, $50+13.65=63.65$. Por

lo tanto, $P_{95} = 63.65$. Su interpretación indicaría que el 95% de las mujeres que tuvieron Covid-19 de tipo relacionado, en estado leve y que se han recuperado presentaron una edad máxima de 63.65 años, aproximadamente 64 años.

OBJETIVOS

Utilizar la aplicación Microsoft Excel para el cálculo de medidas de posición: cuartiles, deciles y percentiles en ejercicios de contexto, contemplando datos desagrupados en variable cuantitativa continua y discreta.

MATERIALES

- Equipos de cómputo
- Microsoft Excel

DESCRIPCIÓN DE ACTIVIDADES Y PROCEDIMIENTOS DE LA PRÁCTICA EN MICROSOFT EXCEL

Para realizar el procedimiento en Microsoft Excel, se procederá a realizar los ejercicios de variable cuantitativa continua (el procedimiento para variable cuantitativa discreta es igual) resueltos en

la sección teórica de medidas de posición: cuartiles, deciles y percentiles, con el fin de verificar los resultados obtenidos en ese apartado.

Ejercicio variable cuantitativa continua

La talla normal de recién nacidos a término en un hospital durante una mañana fue: 46, 48, 46, 50, 53, 51, 47, 47, 51, 52, 50, 48, 49, 50, 51. Determine el cuartil 2, el decil 4 y el percentil 85.

Cuartil 2 (Q_2)

Abrimos Microsoft Excel, creamos un nuevo libro y nos ubicamos en una hoja en la que registraremos los datos. No es necesario digitar los datos en orden, se pueden digitar en el orden que se encuentran en el contexto planteado en la figura 2.

Para el cálculo de los cuartiles se utilizará la función Cuartil.exc. Esto es, nos ubicamos en una celda en que se registrará el valor resultante del cuartil, en este caso, la celda B4, se inserta el símbolo igual (=) y se digita la función, que aparecerá en la primera posición, como se evidencia en la figura 3.

FIGURA 2. Inicio procedimiento para el cálculo del cuartil.

Fuente: elaboración propia.

FIGURA 3. Búsqueda de función cuartil.exc.

Fuente: elaboración propia.

Seleccionamos Cuartil.exc, como se muestra en la figura 4.

FIGURA 4. Selección de función cuartil.exc

Fuente: elaboración propia.

La función solicita dos argumentos:

Matriz: obligatorio. Es la matriz o el rango de celdas de valores numéricos cuyo cuartil desea obtener. En este caso, serán la selección desde la celda B2 hasta la celda P2

Cuartil: obligatorio. Indica el valor que se devolverá. Para el ejercicio será 2, ya que el ejercicio solicita el cuartil 2.

FIGURA 5. Selección de datos para uso de función Cuartil.exc.

Fuente: elaboración propia.

Luego damos *Enter*, obteniendo lo que se muestra en la figura 6.

FIGURA 6. Resultado del uso de la función cuartil.exc.

Fuente: elaboración propia.

Es decir, el resultado da 50. Esto concuerda con el procedimiento realizado en el ejercicio resuelto en la sección que explica los cuartiles. Para la interpretación de este valor, diríjase a la sección teórica en el apartado de cuartiles - ejemplo datos cuantitativos continuos.

Diagrama de cajas

Para realizar el gráfico de diagrama de caja de maneja sencilla, seleccionamos los datos que, en este caso, son las tallas de los recién nacidos y damos clic en Insertar como aparece en la figura 7.

FIGURA 7. Opción Insertar de Microsoft Excel.

Fuente: elaboración propia.

Se da clic en la opción Gráficos recomendados y debe mostrar que se muestra en la figura 8.

FIGURA 8. Selección opción Gráficos recomendados.

Fuente: elaboración propia.

Posteriormente dar clic en la opción Todos los gráficos (figura 9).

FIGURA 9. Selección opción todos los gráficos.

Fuente: elaboración propia.

Se busca la opción Caja y bigotes, y se selecciona (figura 10).

FIGURA 10. Selección opción Caja y bigotes.

Fuente: elaboración propia.

Se da Aceptar y aparece el diagrama como en la figura 11.

FIGURA 11. Diagrama de caja del ejercicio de tallas de recién nacidos.

Fuente: elaboración propia.

Al dar clic en el más (+), se va a deseleccionar Título del gráfico y seleccionar Etiqueta de datos, de la manera como se muestra en la figura 12.

FIGURA 12. Ajustes del diagrama de caja.

Fuente: elaboración propia.

De esta manera, se obtiene el diagrama de caja correspondiente al ejemplo de recién nacidos, en que se evidencian los valores establecidos en la figura 1. El valor que se ubica dentro de la caja y con cifras decimales es la media aritmética simple de todos los 15 datos de tallas de los recién nacidos.

Decil 4 (D_4):

Continuaremos con la solución del ejercicio, pero ahora determinaremos el decil 4 (figura 13).

FIGURA 13. Inicio para procedimiento para el cálculo del decil.

Fuente: elaboración propia.

Para el cálculo de los deciles se utilizará la función `Percentil.exc`, es decir, Microsoft Excel contempla función para cuartiles y percentiles. Sin embargo, la función `Percentil.exc` puede ser aplicada no solo para hallar los percentiles, sino también los deciles, por ende, nos ubicamos en una celda donde se registrara el valor resultante del decil, en este caso, la celda B5, se coloca el símbolo igual (=) y se digita la función, que aparecerá en la primera posición, como se evidencia en la figura 14.

La función solicita dos argumentos:

Matriz: obligatorio. Es la matriz o el rango de celdas de valores numéricos cuyo decil desea obtener. En este caso, serán la selección desde la celda B2 hasta la celda P2.

k: obligatorio. Es el valor de percentil en el rango de 0 a 1, ambos incluidos. Para el desarrollo del ejercicio, será 0.4, debido a que el decil que se desea determinar es el decil 4 que equivaldría al 40%.

FIGURA 14. Búsqueda de función Percentil.exc para el cálculo del decil.

Fuente: elaboración propia.

Seleccionamos Percentil.exc y queda como se muestra en la figura 15.

FIGURA 15. Selección de la función Percentil.exc para el cálculo del decil.

Fuente: elaboración propia.

FIGURA 16. Selección de datos para el uso de la función Percentil.exc para el cálculo del decil.

Fuente: elaboración propia.

Luego damos *Enter*, obteniendo lo que se evidencia en la figura 17.

FIGURA 17. Resultado del uso de la función percentil.exc para el cálculo del decil.

Fuente: elaboración propia.

Es decir, el resultado es 48.4, lo cual concuerda con el procedimiento realizado del ejercicio resuelto en la sección teórica de la explicación de deciles. Para la interpretación de este valor, dirijase a la sección teórica en el apartado de deciles (*ejemplo datos cuantitativos continuos*).

Percentil 85 (P_{85})

Finalizaremos el ejercicio determinando el percentil 85.

FIGURA 18. Inicio procedimiento para el cálculo del percentil.

Fuente: elaboración propia.

Para el cálculo del percentil se utilizará nuevamente la función `Percentil.exc`, por tanto, nos ubicamos en una celda en que se registrará el valor resultante del decil, en este caso, la celda B6, se coloca el símbolo igual (=) y se digita la función, que aparecerá en la primera posición, como se evidencia en la figura 19.

La función solicita dos argumentos:

Matriz: obligatorio. Es la matriz o el rango de celdas de valores numéricos cuyo decil desea obtener. En este caso, serán la selección desde la celda B2 hasta la celda P2.

k: obligatorio. Es el valor de percentil en el rango de 0 a 1, ambos incluidos. Para el desarrollo del ejercicio, será 0.85, debido a que el decil que se desea determinar es el percentil 85 que equivaldría al 85%.

FIGURA 19. Búsqueda de la función Percentil.exc.

Fuente: elaboración propia.

Seleccionamos Percentil.exc, tal como se muestra en las figuras 20 y 21.

FIGURA 20. Selección de la función Percentil.exc.

Fuente: elaboración propia.

FIGURA 21. Selección de los datos para el uso de la función Percentil.exc.

Fuente: elaboración propia.

Luego damos Enter, obteniendo lo que aparece en la figura 22.

FIGURA 22. Resultado del uso de la función Percentil.exc.

Fuente: elaboración propia.

Es decir, el resultado es 51.6, lo que concuerda con el procedimiento realizado en el ejercicio resuelto en la sección teórica de la explicación

de percentiles. Para la interpretación de este valor, dirijase a la sección teórica en el aparte de percentiles (*ejemplo datos cuantitativos continuos*).

TRABAJO PARA LOS ESTUDIANTES

1. En un examen de ingreso a la Universidad se aplicó una prueba sobre comprensión lectora. Los resultados obtenidos por los estudiantes que aspiraban al programa de Odontología fueron los siguientes:

Aspirantes a Odontología

3,0	3,7	2,5	2,2	3,5	3,0
2,7	3,8	3,8	2,2	2,5	3,8
3,3	2,8	2,7	3,5	2,2	2,7
2,2	2,3	2,3	2,3	2,0	3,8
2,7	2,2	3,2	2,5	3,7	2,3
3,5	2,2	2,0	2,5	3,0	2,8
3,8	2,3	3,0	2,0	2,3	3,7
2,7	2,2	2,3	2,7	2,7	2,2
3,3	2,0	3,0	2,2	2,3	2,3
3,5	2,8	3,0	3,0	2,2	4,0

Determine el cuartil 2, el decil 6, el decil 3, el percentil 80, el percentil 48 y realice el diagrama de caja. Desarrolle el procedimiento como se explicó en los ejercicios resueltos en la sección teórica y luego valide los resultados mediante Microsoft Excel (Hernández, 2016).

2. Considere las siguientes estaturas (en cm) de ciertos trabajadores en una empresa.

Altura:			172	162	147	184	140	156
153	186	157	189	162	175	162	158	163
150	152	163	151	182	146	154	163	170
183	162	176	167	168	165			

Determine el cuartil 1, el decil 5, el decil 7, el percentil 60, el percentil 75 y realizar el diagrama de caja. Desarrolle el procedimiento como se explicó en los ejercicios resueltos en la sección teórica y luego valide los resultados mediante Microsoft Excel (CIDEAD, 2020).

3. Se solicitó a 30 empleados de una institución que participaran en una prueba que mide la capacidad para el ejercicio. Se midió esta capacidad por empleado (en minutos):

25	27	30	33	30	32	30	34	30	27	
26	25	29	31	31	32	34	32	33	30	27
	30	31	36	28	30	31	26	29	32	

Determine el cuartil 2, el decil 3, el decil 8, el percentil 50, el percentil 85 y realice el diagrama de caja. Desarrolle el procedimiento como se explicó en los ejercicios resueltos en la sección teórica y luego valide los resultados mediante Microsoft Excel (Muñoz, 2020).

REFERENCIAS

- CIDEAD. (25 de julio de 2020). CIDEAD. Recuperado de: CIDEAD: http://recursostic.educacion.es/secundaria/edad/3esomatematicas/3quincena11/index3_11.htm
- Hernández, G. J. (2016). *Elementos Básicos de estadística descriptiva para el análisis de datos*. Medellín: Fondo Editorial Luis Amigó.
- Instituto Nacional de Salud (26 de julio de 2020). Boletines casos Covid-19 Colombia. Recuperado de: <https://www.ins.gov.co/Paginas/Boletines-casos-COVID-19-Colombia.aspx>
- Lind, D. A., Marchal, W., y Wathen, S. (2015). *Estadística aplicada a los negocios y la economía*. México: McGrawHill.
- Martín, Á. T. (2016). El uso de la herramienta “análisis de datos” de Excel como complemento para la enseñanza de la estadística en el aula. *Congreso Uruguayo de Educación Matemática*. Chile: CUREM6.
- Martínez, C. (2012). *Estadística y muestreo*. Bogotá: ECOE.
- Martínez, C. (2016). *Estadística básica aplicada*. Bogotá: Ediciones ECOE.
- Muñoz, P. J. (25 de julio de 2020). Curso de estadística básica. Recuperado de: <https://es.slideshare.net/pablo-jesuscontrerasmunoz/estadistica-basica-sesion-3>