MATEMÁTICA

Medidas de Tendencia Central

****

**Liceo Técnico Profesional José Miguel Carrera**

**José Francisco Vergara #399, Quilicura**

Felipe Morales M.

# Medidas de Tendencia Central

En Chile existen numerosas instituciones que realizan investigaciones estadísticas sobre diferentes temas, tales como la salud pública, la educación, la cultura, la locomoción colectiva, entre otros.

 En tu caso cuando en el colegio te entregan una nota 5,0 en una prueba, tiendes a preguntar cómo les fue a tus amigos para así saber que tan variadas fueron las notas del curso, esto te permite saber que tan buena o mala fue tu nota 5,0 ya que la nota por si sola tiende a carecer de significado. En situaciones como estas, de forma intuitiva, tiendes a realizar un estudio sobre las notas de tu curso de forma tal que puedas decir frases como “me fue súper bien en la prueba, estuve por sobre el promedio de mi curso”, “la prueba estuvo difícil, fui la nota más baja” o “soy del montón, la mayoría de mis compañeros obtuvo nota 5,0”.

 De acuerdo a lo planteado anteriormente, es que nos damos cuenta que no basta con recolectar, organizar y presentar la información en una tabla estadística o en un gráfico, sino que debemos contar con otros elementos de referencia que nos permitan analizar la información desde otra perspectiva. En este caso acudiremos a las medidas de tendencia central para interpretar el comportamiento de los datos.

|  |
| --- |
| Las medidas de tendencia central son valores numéricos que expresan el grado de centralización de los datos que representan. |

Las medidas de tendencia central más utilizadas son:

## Media Aritmética (X)

La media aritmética, también denominada promedio, es una medida de tendencia central que solo se puede aplicar en variables cuantitativas.

|  |
| --- |
|  La media se define como la suma de los valores de todas las observaciones divididos por el número total de datos.  |

Algunas ideas sobre esta medida de tendencia central son:

* No es necesario que los datos estén ordenados para calcular la media aritmética.
* Todos los datos son incluidos en el cálculo de la media aritmética.
* Un conjunto de datos solo tiene una media aritmética.
* El valor numérico puede o no coincidir con algunos de los datos del conjunto.
* Se utiliza generalmente para comparar dos o más conjuntos de datos.
* Es sensible a una distribución muy asimétrica de los datos, es decir, pierde precisión cuando hay valores extremos, muy grandes o muy pequeños, en comparación con el general de la muestra.
* Cuando se aplica en datos agrupados en intervalos, la medida pierde precisión debido a que existe una pérdida de información al agrupar los datos en clases.

A continuación mostraremos como calcular la media aritmética en distintas situaciones de acuerdo a como se nos presentan los datos:

1. *Con miras a las compras previas a las fiestas de fin de año, el Servicio Nacional del Consumidor (SERNAC) decide realizar un sondeo para conocer el costo de una cena familiar. La muestra se tomó entre el 11 y 13 de Diciembre del 2012 y los resultados que arrojo en cuanto al precio del producto “Duraznos mitades, grado 2, Dos Caballos” en diferentes sectores de Santiago se muestran a continuación:*

$849 − $856 − $889 − $854 − $907

En esta situación se nos presentan los datos por extensión, por lo tanto, basta con sumar cada dato y dividirlo por el total de estos para obtener la media aritmética.

= $\frac{849 + 856 + 889 + 854 + 907 }{5}$

=871

En base al desarrollo anterior, podemos decir que el valor promedio del producto “Duraznos mitades, grado 2, Dos Caballos” es de **$871**.

De forma general, cuando se nos presentan los **datos por extensión**, el método para encontrar la media aritmética consiste en:

*a) Sumar todos los datos (xi).*

*b) Dividir el resultado de la suma en el total de datos (n).*

De esta forma la media aritmética es:

= $\frac{x1 + x2 + x3 + . . . + xk }{n}$

1. *Una organización que promueve la vida saludable decide estudiar los kilogramos de fruta que compran 120 familias. Durante una semana se registraron los kilogramos de fruta que compraron cada familia, obteniéndose los siguientes datos tabulados:*

|  |  |
| --- | --- |
| KILOS | FAMILIAS |
| 1 | 5 |
| 2 | 18 |
| 3 | 26 |
| 4 | 17 |
| 5 | 12 |
| 6 | 34 |
| 7 | 8 |

En esta situación nos entregan los datos tabulados con su respectiva frecuencia, por lo tanto en vez de sumar 5 veces el número uno, acudiremos a la multiplicación, de esta forma en vez de operar 1 + 1 + 1 + 1 + 1 reduciremos la expresión a 1 · 5. Al realizar lo mismo con todos los datos de la tabla obtenemos la media aritmética de la siguiente forma:

= $\frac{1 · 5 + 2 · 18 + 3 · 26 + 4 · 17 + 5 · 12 + 6 · 34 + 7 · 8 }{5+ 18 + 26 + 17 + 12 + 34 + 8}$

= $\frac{ 507}{120}$

=4,225

A partir del resultado anterior podemos decir que las familias consumen 4, 225 kilogramos de fruta en promedio a la semana.

 De forma general, en una **tabla de distribución de frecuencia**, el método para encontrar la media aritmética consiste en:

*a) Multiplicar cada dato (xi) por su frecuencia (fi).*

*b) Sumar todos los resultados anteriores.*

*c) Dividir el resultado de la suma en el total de datos (n). De esta forma la media aritmética es:*

= $\frac{x1 · f1 + x2 · f2 + . . . + xk · fk }{n}$

1. *Para que un estudiante pueda entrar a la Universidad de Chile a estudiar Ingeniería y Ciencias-Plan Común, debe postular con un puntaje PSU que se obtiene de acuerdo al valor que tiene cada puntaje para la carrera. A continuación se muestran los puntajes obtenidos por un estudiante en el proceso de admisión 2013:*

|  |  |  |
| --- | --- | --- |
| REQUISITOS | PUNTAJES | PONDERACIÓN (%) |
| Notas de Enseñanza Media | 723 | 10 |
| Ranking | 770 | 10 |
| Prueba de Lenguaje | 665 | 10 |
| Prueba de Matemática | 746 | 50 |
| Prueba de Ciencia | 708 | 20 |

En este caso para obtener el puntaje con el cual el estudiante postula a la carrera universitaria se debe calcular la media aritmética ponderara ya que no todos los datos tienen igual valor.

= $\frac{723 · 10 + 770 · 10 + 655 · 10 + 746 · 50 + 708 · 20 }{10 + 10 + 10 + 50 + 20}$

= $\frac{72.940 }{100}$

=729,4

En base a lo anterior, el estudiante postula con 729,4 puntos al Plan Común de Ingeniería y Ciencias de la Universidad de Chile.

De forma general, en una tabla con datos ponderados, el método para encontrar la media aritmética consiste en:

*a) Multiplicar la ponderación (pi) por su dato (xi).*

 *b) Sumar todos los resultados anteriores.*

*c) Dividir el resultado de la suma total de las ponderaciones (P).*

 De esta forma la media aritmética es: X = x1 · p1 + x2 · p2 + . . . + xk · pk P

1. *En un hospital público se desea estudiar el tiempo de espera de los pacientes entre las 22:00 horas y las 00:00 horas. Durante un día viernes se registraron los tiempos de espera de los pacientes, obteniéndose los siguientes datos tabulados:*

|  |  |  |
| --- | --- | --- |
| Tiempo [min] | Pacientes | Marca de Clase |
| [0-30[ | 5 | 15 |
| [30-60[ | 8 | 45 |
| [60-90[ | 6 | 75 |
| [90-120[ | 15 | 105 |
| [120-150[ | 35 | 135 |
| [150-180[ | 12 | 165 |
| [180-210[ | 20 | 195 |
| [210-240[ | 24 | 225 |
| [240-270[ | 50 | 255 |
| [270-300[ | 40 | 285 |

En esta situación no conocemos los datos recolectados por el hospital sino que solo conocemos los intervalos en los que estos están agrupados, por lo tanto, para calcular la media aritmética haremos uso de la marca de clase[[1]](#footnote-2)debido a que este es un valor representativo de cada intervalo.

= $\frac{5·15 + 8·45 + 6·75 + 15·105 + 35·135 + 12·165 + 20·195 + 24·225 + 50·255 + 40·285}{5 + 8 + 6 + 15 + 35 + 12 + 20 + 24 + 50 + 40}$

=$\frac{75+ 360 + 450+1.575+4.725 + 1.980 + 3.900 + 5.400 + 12.750 + 11.400}{215}$

= $\frac{42.615 }{215}$

≈198,21

De acuerdo al cálculo anterior, el tiempo de espera promedio por los pacientes del hospital público fue de 198,21 minutos.

De forma general, en una **tabla de frecuenciacon datos agrupados en intervalos**, el método para encontrar la media aritmética consiste en:

a*) Multiplicar la marca de clase (mi) de cada intervalo por su frecuencia (fi).*

*b) Sumar todos los resultados anteriores.*

*c) Dividir el resultado de la suma en el total de datos (n).*

De esta forma la media aritmética es:

= $\frac{m1 · f1 + m2 · f2 + · · · + mk · fk}{n}$

## Mediana (Me)

La mediana es una medida de tendencia central que es aplicada solo en variables cuantitativas.

|  |
| --- |
| La mediana se define como el valor numérico que divide a un conjunto de datos, ordenados de manera creciente o decreciente, en dos partes iguales, es decir, deja por debajo y por encima de sı el 50 % de la distribución de datos. |

Algunas ideas sobre esta medida de tendencia central son:

* Es necesario que los datos estén ordenados para calcular la mediana.
* Un conjunto de datos solo tiene una mediana.
* El valor numérico puede o no coincidir con algunos de los datos del conjunto.
* Es estable a los valores extremos de un conjunto de datos.

 A continuación mostraremos como calcular la mediana en distintas situaciones de acuerdo a como se nos presentan los datos:

1. *El Servicio Nacional del Consumidor (SERNAC) decide realizar un sondeo para conocer el precio del combustible lıquido según sector de la Región Metropolitana. El estudio se llevó a cabo el día 31 de Diciembre del 2012 y los resultados registrados sobre el precio promedio de la gasolina de 97 octanos se presentan a continuación:*

$780 − $774 − $792 − $771 − $776

 En este caso el número de datos es impar, por lo tanto, para calcular la mediana basta con ordenar los datos de forma creciente o decreciente y determinar el valor central.

$771 − $774 − $776 − $780 − $792

De acuerdo al desarrollo anterior, la mediana del precio de la gasolina de 97 octanos es $776, valor que corresponde al tercer lugar (X3) de la distribución ordenada de datos.

De forma general, cuando se nos presenta un conjunto impar de datos discretos por extensión, el método para encontrar la mediana consiste en:

*a) Ordenar los datos de forma creciente o decreciente.*

*b) Localizar el valor que divide en dos partes iguales al total de datos (n).*

De esta forma la mediana es el dato que ocupa el lugar:

Me= $\frac{X(n+1)}{2}$

1. *El Servicio Nacional del Consumidor (SERNAC) realizo un sondeo sobre el precio del pan durante el año 2012 en la Región Metropolitana. Los registros tomados en el mes de Septiembre para 6 tipos de panes se muestran a continuación:*

$962 − $912 − $1.239 − $1.174 − $1.342 − $1.325

En esta situación el número de datos es par, por lo tanto, para calcular la mediana tendremos que calcular el promedio entre los dos datos centrales que tengamos luego de ordenar nuestra información de forma creciente o decreciente.

$912 − $962 − $1.174 − $1.239 − $1.325 − $1.342 7

Me= $\frac{1.174 + 1.239}{2}$

Me= $\frac{2.413}{2}$

Me = 1.206, 5

De acuerdo al desarrollo anterior, la mediana del precio del pan en la Región Metropolitana es de $1.206,5.

De forma general, cuando se nos presenta un conjunto par de datos discretos por extensión, el método para encontrar la mediana consiste en:

*a) Ordenar los datos de forma creciente o decreciente.*

*b) Localizar los dos valores centrales (En/2 y X(n/2)+1) de la distribución total de datos (n). c) Calcular el promedio entre los dos valores encontrados anteriormente.*

De esta forma la mediana es:

Me= $\frac{Xn/2 +X(n/2)+1 }{2}$

1. *El 15 de Diciembre se realizó la corrida Nike “We run Santiago 10K 2012”. El tiempo que se demoraron en recorrer los primeros 5 kilómetros 72 mujeres entre 16 años y 19 años se encuentra registrado en la siguiente tabla:*

|  |  |
| --- | --- |
| T [min] | Mujeres |
| [20-25[ | 1 |
| [25-30[ | 8 |
| [30-35[ | 23 |
| [35-40[ | 30 |
| [40-45[ | 6 |
| [45-50[ | 4 |

En una tabla de frecuencia con datos agrupados en intervalos, el método para encontrar la mediana consiste en:

*a) Determinar el valor numérico de la mitad de los datos (n/2)*

*b) Localizar el intervalo en el cual está contenido ese valor.*

*c) Sustituir los siguientes valores:*

- **n** =Numero de observaciones.

- **a** =Amplitud del intervalo seleccionado.

- **Li** =Lımite inferior del intervalo seleccionado.

- **fi** =Frecuencia absoluta del intervalo seleccionado.

- **Fi** =Frecuencia absoluta acumulada del intervalo anterior a la clase que contiene a la mediana.

En la expresión:

$$Me=Li+a·\left[\frac{\frac{n}{2} - Fi}{fi}\right]$$

En esta situación tenemos un total de 72 mujeres. Para saber en qué intervalo esta nuestra mediana dividimos el total de datos en 2:

$$\frac{72}{2}=36$$

De acuerdo a lo anterior, nuestra mediana corresponde al dato número 36, que está ubicado de la clase [35 − 40[.

Luego de tener nuestro intervalo identificado, determinamos los valores numéricos de los datos necesarios para obtener la mediana:

- n = 72

- a = 40 − 35 = 5

- Li = 35

- fi = 30

- Fi = 1 + 8 + 23 = 32

Finalmente sustituimos en la expresión:

$$Me=35+5·\left[\frac{\frac{n}{2} - Fi}{fi}\right]$$

$$Me=35+5·\left[\frac{\frac{72}{2} - 32}{30}\right]$$

$$Me=35+5·\left[\frac{36 - 32}{30}\right]$$

$$Me=35+5·\left[\frac{4}{30}\right]$$

$$Me=35+\left[\frac{2}{3}\right]$$

$$Me=\left[\frac{107}{3}\right]$$

$$Me ≈35,67$$

La mediana de las mujeres que corrieron el evento “We run Santiago 10K 2012” es de 35,67 minutos. 9 open green road

## 1.3. Moda (Mo)

La moda es una medida de tendencia central que es aplicada en variables cuantitativas y variables cualitativas.

|  |
| --- |
| La moda se define como el dato que posee mayor frecuencia absoluta, es decir, el valor que más se repite.  |

* Algunas ideas sobre esta medida de tendencia central son:
* No es necesario que los datos estén ordenados para calcular la moda.
* Un conjunto de datos puede tener más de una moda o puede que esté presente.
* El valor numérico coincide con algún dato del conjunto.

A continuación mostraremos como calcular la moda en distintas situaciones de acuerdo a como se nos presentan los datos:

1. *El Servicio Nacional del Consumidor (SERNAC) decide realizar el sondeo mensual de precios de los medicamentos en farmacias del área metropolitana correspondiente a Noviembre del 2012. Los precios del medicamento “Clorfenamina Maleato, 4 mg de 20 Grageas, Laboratorio Chile” registrados en distintas farmacias se muestran a continuación:*

$190 − $195 − $290 − $290 − $350 − $240

En este caso para encontrar la moda debemos determinar el valor que más se repite dentro de los datos.

$190 − $195 − $290 − $290 − $350 − $240

En este caso la moda es $290 con una frecuencia de 2.

1. *El 15 de Diciembre se realizó la corrida Nike “We run Santiago 10K 2012”. El tiempo que se demoraron en realizar la corrida los 30 participantes hombres entre 60 años y 64 años se encuentran registrados en la siguiente tabla:*

|  |  |
| --- | --- |
| Tiempo [min] | Hombres |
| [40-50[ | 2 |
| [50-60[ | 10 |
| [60-70[ | 13 |
| [70-80[ | 4 |
| [80-90[ | 1 |

Al presentarnos la información a través de una tabla de frecuencia con datos agrupados en intervalos, se acepta como valido que la moda corresponda a la marca de clase del intervalo que posea mayor frecuencia absoluta.

En esta situación, el intervalo que tiene mayor frecuencia absoluta corresponde a [60, 70[ con 13 adultos que realizaron la carrera dentro de ese tiempo. De acuerdo a lo descrito arriba la moda serıa:

$$Mo=\left[\frac{60+70}{2}\right]$$

$$Mo=\left[\frac{130}{2}\right]$$

$$Mo=\left[65\right]$$

En este caso la moda es de 65 minutos. Sin embargo, cuando se desea mayor precisión se puede trabajar de la siguiente forma:

*a) Localizar el intervalo que posee mayor frecuencia absoluta.*

*b) Sustituir los siguientes valores:*

- a =Amplitud del intervalo seleccionado.

- Li =Lımite inferior del intervalo seleccionado.

- fi =Frecuencia absoluta del intervalo seleccionado.

- fi+1 =Frecuencia absoluta de la clase siguiente al intervalo seleccionado.

 - fi−1 =Frecuencia absoluta de la clase anterior al intervalo seleccionado.

En la expresión:

$$Mo=Li+a·\left[\frac{fi -(fi-1)}{fi-\left(fi-1\right)+fi-(fi+1)}\right]$$

En esta situación el intervalo con mayor frecuencia es [60 − 70[, por lo tanto, los datos serıan:

- a = 70 − 60 = 10

- Li = 60

- fi = 13

- fi+1 = 4

- fi−1 = 10

Al sustituir en la expresión antes mencionada tenemos:

$$Mo=Li+a·\left[\frac{fi -(fi-1)}{fi-\left(fi-1\right)+fi-(fi+1)}\right]$$

$$Mo=60+10·\left[\frac{13 -10}{\left(13-10\right)+(13-4)}\right]$$

$$Mo=60+10·\left[\frac{3}{ 3+9}\right]$$

$$Mo=60+10·\left[\frac{3}{ 12}\right]$$

$$Mo=60+10·\left[\frac{ 1}{ 4}\right]$$

$$Mo=60+\left[\frac{ 10}{ 4}\right]$$

$$Mo=60+\left[\frac{5}{ 2}\right]$$

$$Mo≈62,5$$

Por lo tanto, si queremos ser más precisos con el valor de la moda, esta serıa de **62,5 minutos**.

1. la marca de clase corresponde al promedio entre los límites de la clase o intervalo. [↑](#footnote-ref-2)