



# **Estadística básica**

## **Unidad 3. Muestreo, medidas de tendencia central y de dispersión**

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Índice

---

Presentación de la Unidad .....	3
3.1. Muestreo.....	5
3.1.1. Conceptos básicos de muestreo aleatorio.....	5
3.1.2. Metodología del muestreo aleatorio simple .....	6
3.2. Medidas de tendencia central .....	10
3.2.1. Datos no agrupados (media, mediana y moda).....	10
3.2.2. Datos agrupados (media, mediana y moda).....	13
3.3. Medidas de dispersión .....	17
3.3.1. Datos no agrupados (varianza y desviación estándar) .....	17
3.3.2. Datos agrupados (varianza y desviación estándar).....	18
Cierre de la Unidad .....	20
Fuentes de consulta.....	20

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Presentación de la Unidad

Para cualquier conjunto de datos estudiados es importante tener información resumida de sus características, la cual debe indicar cómo se comporta la población de datos que se tiene. Para resumir la información se utilizan dos tipos de valores que, en lugar de representar cada dato, representan conjuntos de datos. Estos dos tipos de indicadores estadísticos son: las medidas de tendencia central, que nos muestran hacia qué valores se agrupan o acumulan los datos, y las medidas de dispersión, que, de forma contraria a las anteriores, muestran cómo se dispersan o separan los datos.

### Propósitos

Mediante el estudio de esta unidad se busca que logres los siguientes propósitos:

- Identificar problemas de datos desagrupados y agrupados.
- Obtener una muestra a partir de datos desagrupados y agrupados.
- Calcular medidas de tendencia central y dispersión.
- Aplicar la estadística básica para elaborar conclusiones.

### Competencia específica

Analizar muestras de una población para explicar fenómenos mediante la interpretación y representación de las medidas de tendencia central y dispersión.



### Para saber más

Se ha seleccionado una serie de recursos en línea con el fin de ofrecerte un panorama general de la unidad y alternativas en caso de que se te dificulte la comprensión de algún concepto o proceso. Estos recursos tienen una extensión breve, pero complementan lo expuesto en este material. Algunos de ellos te ayudarán a visualizar cómo puedes aplicar estos contenidos en las actividades de la unidad. Puedes consultarlos en la sección *Material de apoyo*, en este momento o volver cuando hayas concluido la revisión de esta unidad.

- Alegre J. y Cladera M. (s.f.). *Introducción a la estadística descriptiva para economistas* (p. 11-22, 39-83). Recuperado de <https://www.studocu.com/es-mx/document/universidad-autonoma-del-estado-de-mexico/matematicas/introduccion-a-la-estadistica-descriptiva-para-economistas/7405494>
- Estuardo, A. (2012). *Estadística y probabilidades* (p. 39-51). Recuperado de <https://fddocuments.ec/reader/full/estadistica-y-probabilidades-g-aaron-estuardo-morales>
- Comunidad de Madrid (s.f.). *Estadística básica* (p. 4-13). Recuperado de [http://www.madrid.org/cs/StaticFiles/Emprendedores/Analisis\\_Riesgos/pages/pdf/estadisticas\\_es.pdf](http://www.madrid.org/cs/StaticFiles/Emprendedores/Analisis_Riesgos/pages/pdf/estadisticas_es.pdf)
- Matus J. (s.f.). Fascículo 2. Medidas de tendencia central. En *Estadística descriptiva e inferencial I*. Recuperado de: [https://www.conevyt.org.mx/bachillerato/material\\_bachilleres/cb6/5sempdf/edin1/edin1\\_f1.pdf](https://www.conevyt.org.mx/bachillerato/material_bachilleres/cb6/5sempdf/edin1/edin1_f1.pdf)
- Universidad Autónoma de Querétaro. (s.f.). *Estadística descriptiva* (p. 16-32).

#### Página web

- Universidad de Guanajuato. (s.f.) Programas informáticos para análisis estadístico de datos cuantitativos. <https://blogs.ugto.mx/mdued/wp-content/uploads/sites/66/2022/10/Investigacio%CC%81n-cuantitativa-Ana%CC%81lisis-de-datos-pa%CC%81ginas-35-59.pdf>

### 3.1. Muestreo



Muestra.

Fuente: cooldesign, 2012. [freedigitalphotos.net](http://freedigitalphotos.net)

Los estudios estadísticos normalmente se hacen con una parte de la población, ya que realizarlos sobre la totalidad resultaría demasiado complicado. Para que la información obtenida tenga validez y confiabilidad es necesario que la muestra cumpla con ciertas condiciones específicas, relacionadas con el método para determinar el tamaño y características de la muestra y los individuos que la componen.

#### 3.1.1. Conceptos básicos de muestreo aleatorio

Para que la información obtenida tenga validez y confiabilidad es necesario que cumpla con algunas condiciones específicas. Los métodos de muestreo se pueden clasificar en:

##### Muestreo probabilístico

En él, todos los elementos de una población y, por lo tanto, todas las muestras posibles, tienen la misma posibilidad de ser elegidas. Las muestras obtenidas a través de este tipo de muestreo son confiables porque aseguran la condición de representatividad que es muy importante para hacer generalizaciones.

### Muestreo no probabilístico

En este tipo de muestreo los elementos de la población no comparten las mismas posibilidades de ser seleccionados. Las muestras obtenidas no cumplen con la condición de representatividad, por lo que no es confiable hacer generalizaciones a toda la población.

### 3.1.2. Metodología del muestreo aleatorio simple

Dentro del muestreo probabilístico existen diversos métodos para obtener el tamaño de una muestra, a continuación, estudiarás el muestreo aleatorio simple, el cual consiste en los siguientes pasos.

#### 1. Definir la población de estudio y el parámetro a estudiar

Como recordarás, la **población** es el grupo formado por el **conjunto total** de individuos, objetos o medidas que poseen algunas características comunes observables en un lugar y en un momento determinado. Por lo tanto, el paso 1 es determinar la que se va a estudiar.

Por ejemplo: un investigador realiza un estudio sobre las relaciones de género en el noviazgo, su objeto de estudio son las manifestaciones de violencia física y psicológica entre los estudiantes del último año de la carrera de ingeniería. Su población es el total de estudiantes del último año de ingeniería que tengan novio o novia; el total de individuos con esta característica es de 386 en este ejemplo. Por lo que, la población es de 386 individuos y las variables son: violencia física y violencia psicológica.

#### 2. Enumerar a todas las unidades de análisis que integran la población, asignándoles un número de identidad o identificación

Una vez que se ha definido la población y las variables a estudiar, es necesario asignar un número de identificación a cada individuo de la población.

Siguiendo con el ejemplo de las relaciones de género en el noviazgo en los estudiantes de ingeniería, lo que sigue es enumerar a los 386 estudiantes con un número del 1 al 386.

#### 3. Definir la población de estudio y el parámetro a estudiar

Definir el tamaño de la población significa determinar el número de individuos que la constituyen; la variable  $N$  representa el tamaño de la población. Esto es,  $N=X$ .

Para calcular el tamaño de una muestra hay que tomar en cuenta tres factores:

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

- El porcentaje de confianza con el cual se quiere generalizar los datos desde la muestra hacia la población total.
- El porcentaje de error que se pretende aceptar al momento de hacer la generalización.
- El nivel de variabilidad que se calcula para comprobar la hipótesis.

A continuación, se describen los conceptos enlistados:

### Porcentaje de confianza

Es el grado o nivel de seguridad que existe para generalizar los resultados obtenidos. Esto quiere decir que un porcentaje del 100% equivale a decir que no existe ninguna duda para generalizar tales resultados, pero también implica estudiar a la totalidad de los casos de la población (censo). Se denota como  $Z$ .

Para evitar un costo muy alto se busca un porcentaje de confianza menor, comúnmente es un 95%. El nivel de confianza es la probabilidad que establecemos (sin hacer ningún cálculo) para poder acertar al valor verdadero de la población.

Nota: Al estandarizar este valor, el 95% de confianza corresponde a una  $Z=1.96$ .

### Porcentaje de error

Este error es una distancia alrededor del valor que se desea estimar y da un margen de aproximación. Al igual que en el caso de la confianza, si se quiere eliminar el riesgo del error y considerarlo como 0%, entonces la muestra es del mismo tamaño que la población (censo), por lo que conviene realizar un muestreo que implica menos tiempo y menor costo, aunque se corre un cierto riesgo de equivocarse. Comúnmente se aceptan entre el 4% y el 6% como error, tomando en cuenta que no son complementarios la confianza y el error, es decir, que en un muestreo podemos tener 95% de confianza con 6% de error.

### Variabilidad

Es la probabilidad (o porcentaje) con el que se aceptó y se rechazó la hipótesis que se quiere comprobar. El porcentaje con que se aceptó tal hipótesis se denomina variabilidad positiva y se indica con  $p$  (también llamada probabilidad de éxito), y el porcentaje con el que se rechazó la hipótesis es la variabilidad negativa, identificada con  $q$ , también llamada probabilidad de fracaso, y se obtiene como  $q = 1 - p$ .

Variabilidad positiva  $p$  es la probabilidad de que suceda el evento.

Variabilidad negativa  $q$  es la probabilidad de que no suceda el evento.

Para este curso se considerará siempre  $p = 0.5$ , y por lo tanto  $q = 1 - 0.5 = 0.5$

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### 4. Determinar el tamaño óptimo de muestra para el estudio

Una vez que la población, el porcentaje de confianza, el porcentaje de error y el nivel de variabilidad han sido determinados, se debe calcular el tamaño de la muestra. En este paso se utilizan las siguientes fórmulas, en donde la primera implica que no se conoce el tamaño de la población y la segunda se utiliza cuando sí se conoce el tamaño de la población.

#### Desconocimiento del tamaño de la población

##### Fórmula

$n = \frac{Z^2 p q}{E^2}$	$n$ es el tamaño de la muestra
	$Z$ es el nivel de confianza
	$p$ es la variabilidad positiva
	$q$ es la variabilidad negativa
	$E$ es la precisión o error

#### Ejemplo

En un lote grande de medicinas se desea verificar que la proporción de los ingredientes activos sea el adecuado. Se debe determinar el tamaño de la muestra para un nivel de confianza del 95% con un error del 5%. Si la variabilidad es de  $p = q = 0.5$

Solución: Para que el nivel de confianza sea igual al 95% se tiene que  $P(Z) = 0.95$  si  $Z = 1.96$ .

Debido a que la variabilidad y el error se pueden expresar por medio de porcentajes, en el caso necesario, hay que convertir esos valores a proporciones.

Sustitución: Al sustituir los valores en la fórmula se obtienen los siguientes resultados.

$$n = \frac{Z^2 p q}{E^2} \quad \rightarrow \quad n = \frac{(1.96)^2 (0.5)(0.5)}{(0.05)^2} = \frac{0.9604}{0.0025} = 384.16$$

Es decir, se ocupará una muestra de aproximadamente 384 unidades.

#### Conocimiento del tamaño de la población

##### Fórmula

$n = \frac{Z^2 p q N}{NE^2 + Z^2 p q}$	$n$ es el tamaño de la muestra
	$Z$ es el nivel de confianza
	$p$ es la variabilidad positiva
	$q$ es la variabilidad negativa
	$N$ es el tamaño de la población
	$E$ es la precisión o error

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Ejemplo

En un lote de 25,000 cajas de medicina se desea verificar que la proporción de los ingredientes activos sea el adecuado. Se debe determinar el tamaño de la muestra para un nivel de confianza del 95% con un error del 5%. Si la variabilidad es  $p = q = 0.5$

Solución: Para que el nivel de confianza sea igual al 95% se tiene que  $p(Z) = 0.95$  si  $Z = 1.96$

Sustitución: Al sustituir los valores en la fórmula se obtienen los siguientes resultados.

$$n = \frac{Z^2 p q N}{NE^2 + Z^2 p q} \quad n = \frac{(1.96)^2 (0.5)(0.5)(25000)}{(25000 \times 0.05^2) + (1.96^2 \times 0.5 \times 0.5)} = \frac{24010}{62.5 + 0.9604} = 378.34$$

En otras palabras, se ocupará una muestra de aproximadamente 378 cajas.

### 5. Seleccionar la muestra usando números aleatorios

El último paso para obtener la muestra es saber qué individuos específicos de la población se tomarán. Para hacer esto se debe:

1. Numerar a los individuos de la población del 1 al  $N$  (donde  $N$  es el tamaño de la población).
2. Generar números aleatorios mediante herramientas informáticas (por ejemplo, hojas de cálculo con la función “=aleatorio ()”), funciones en calculadora o bien utilizando tablas de números aleatorios. También puedes generar números aleatorios de formas mecánicas, por ejemplo, sacando números de una urna o lanzando una moneda al aire.
3. Tomar los individuos correspondientes a los números elegidos.

Revisa el siguiente recurso para conocer tres formas de obtener números aleatorios.

Ingresa a la sección *Material de apoyo* de la asignatura, para descargar el documento *Obtención de números aleatorios*.

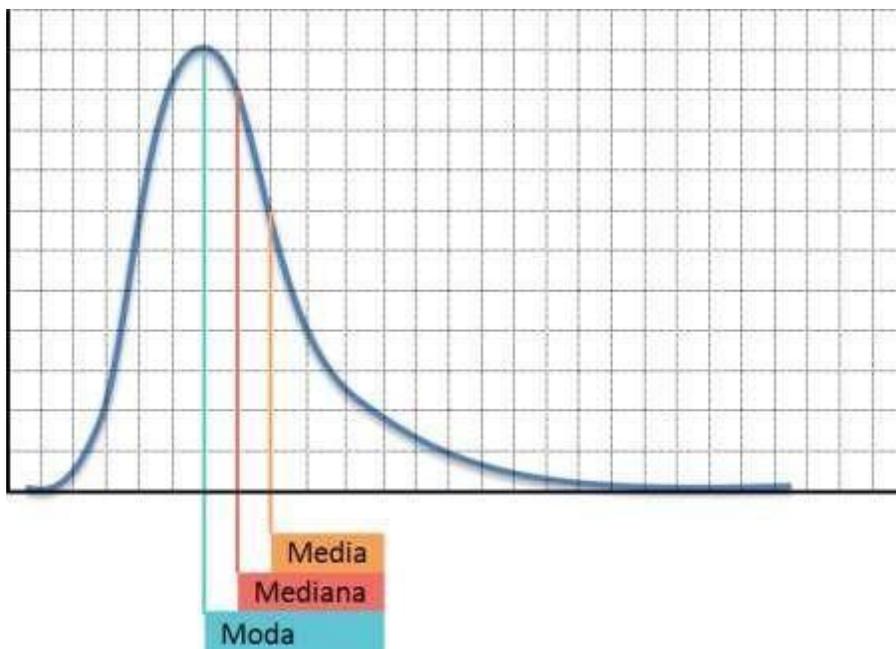


# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### 3.2. Medidas de tendencia central

Las medidas de tendencia central son los valores que representan un conjunto de datos de forma tal que ayudan a saber *dónde* están acumulados los datos, pero sin indicar cómo se distribuyen. Se llaman así porque tienden a ubicarse en la parte central del conjunto de datos. Las medidas de tendencia central más comunes son: la media aritmética, comúnmente conocida como **media o promedio**, la **mediana** y la **moda**.



#### 3.2.1. Datos no agrupados (media, mediana y moda)

Con la finalidad de que las medidas de tendencia central tengan mayor validez estadística se utilizarán fórmulas diferentes para datos agrupados y datos no agrupados, en donde también se deben distinguir si se trabaja con una muestra o con una población.

##### Media

##### Concepto y fórmula

La media aritmética o, simplemente, media, se denota por  $\bar{x}$  o por la letra  $\mu$  según se calcule en una muestra o en la población, respectivamente. La media es el resultado de dividir la suma de todos los valores ( $x_i$ ) entre el número total de datos,  $N$  para el caso de toda la población y  $n$  para el caso de una muestra.

La fórmula para calcular la media de una distribución de datos varía de acuerdo a la manera como se tienen organizados.

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

Fórmula para calcular la media en datos no agrupados: Los datos no agrupados son aquéllos que se organizan en una tabla de datos, es decir, cada valor se representa de manera individual. Las fórmulas para calcular la media son:

En una población	En una muestra
$\mu = \frac{\sum_{i=1}^N X_i}{N}$	$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$

En estas fórmulas la diferencia radica en que el total de la población se representa con la letra  $N$  y el total de la muestra con la letra  $n$ , en donde la media poblacional se denota con la letra griega “Mu” y la media muestral se presenta como “equis barra” .

### Ejemplo

En una serie de días elegidos al azar, se registró el tiempo, en horas, de utilización de dos impresoras en una empresa y se obtuvieron los siguientes resultados:

Impresora I: 3.2, 2.1, 2.7, 3.4, 1.9, 4.2, 3.8, 2.6, 5.2, 4

Impresora II: 3.4, 3.3, 2.5, 4.6, 2.8, 3.6, 4.3

Se requiere lo siguiente: hallar el tiempo medio de utilización de cada impresora.

	Impresora I	Impresora II
1	1,9	2,5
2	2,1	2,8
3	2,6	3,3
4	2,7	3,4
5	3,2	3,6
6	3,4	4,3
7	3,8	4,6
8	4	
9	4,2	
10	5,2	

Datos ordenados

### Respuestas

Para obtener la media de la impresora 1 se suma cada uno de los valores:  $1,9+2,1+2,6+2,7, +3,2+3,4+3,8+4,0+4,2+5,2$  a continuación el resultado de la sumatoria que es 33.1, se divide entre el número de observaciones de la muestra que es 10 y se obtiene el resultado que es 3,31. Análogamente se realiza el mismo procedimiento para la impresora 2 y se obtiene el resultado de 3,5

Media impresora I    3,31  
Media impresora II    3,5

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Mediana

#### Concepto

La mediana (Me) es el valor que divide a la mitad la serie de datos que se tienen. Es decir, la mediana queda en medio de todos los datos cuando los acomodas ya sea en orden creciente o decreciente, entonces, el número de datos que queda a la izquierda de la mediana es igual al número de datos que queda a la derecha.

Si  $n$  es impar hay un dato que queda en medio de todos, éste será igual a la mediana. Si  $n$  es par hay dos datos que quedan en medio de todos, en este caso la mediana es el promedio de esos dos datos, es decir, su suma dividida entre dos.

#### Ejemplos

Para cuando la cantidad de valores de la distribución es impar

Supón que se tienen los siguientes valores: 2, 4, 0, 8, 6, 4, 7, 1, 1, 0, 8, 6, 9

1. Se ordenan los valores de menor a mayor.	0, 0, 1, 1, 2, 4, 4, 6, 6, 7, 8, 8, 9
2. Se busca el valor del centro.	El dato que divide a la mitad es: 4, por lo tanto, la mediana Me: 4

Para cuando la cantidad de valores es par

Supón que se tienen los siguientes valores: 5, 7, 2, 3, 1, 6, 9, 8, 6, 4, 7, 1, 3, 2

1. Se ordenan los valores de menor a mayor.	1, 1, 2, 2, 3, 3, 4, 5, 6, 6, 7, 7, 8, 9
2. SE buscan los valores del centro.	1, 1, 2, 2, 3, 3, 4, 5, 6, 6, 7, 7, 8, 9
3. Se promedian los valores del centro.	$\frac{4+5}{2} = \frac{9}{2} = 4.5$ , por lo tanto Me: 4.5

### Moda

Para el caso de la moda (Mo), en los datos no agrupados, la moda corresponde al valor que más se repite, si se tienen los siguientes datos: 1,1,2,2,3,3,4,4,4,4,5,5,6,6,7,8,9,9,9 la Moda es: 4.

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### 3.2.2. Datos agrupados (media, mediana y moda)

#### Media

##### Fórmulas

###### Fórmula para calcular la media en datos agrupados por frecuencias simples

Los datos agrupados en frecuencias son aquellos que se organizan en una tabla de frecuencias, es decir, las tablas que contienen, en una columna, el valor de la variable ( $x_i$ ) y, en otra columna, la frecuencia ( $f_i$ ) o el número de veces que se repite cada valor en una serie de datos. Para calcular la media con datos agrupados se procede a realizar la sumatoria de el valor de la variable ( $x_i$ ) por el valor de su frecuencia ( $f_i$ ) y el resultado se divide, para el caso de la población, entre  $N$ , y para el caso de la muestra, entre  $n$ .

Las fórmulas para calcular la media con los datos organizados de esta manera son:

En una población	En una muestra
$\mu = \frac{\sum_{i=1}^N x_i f_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^n x_i f_i}{n}$

###### Fórmula para calcular la media en datos agrupados por intervalos

Los datos agrupados en intervalos son aquellos que se organizan dentro de un rango establecido entre un límite inferior y un límite superior. Recuerda que las tablas de intervalos muestran el número de datos que abarca cada intervalo (frecuencia por intervalo).

Las fórmulas para calcular la media con los datos organizados de esta manera son:

En una población	En una muestra
$\mu = \frac{\sum_{i=1}^N Mc_i f_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^n Mc_i f_i}{n}$

En donde debes realizar la sumatoria de cada marca de clase ( $Mc_i$ ) por su frecuencia ( $f_i$ ) y el resultado se divide entre el total de elementos poblacionales ( $N$ ) -si se trata de población- o bien, entre los elementos de la muestra ( $n$ ).

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Ejemplo

En el siguiente cuadro estadístico se presentan, mediante una distribución de frecuencias, los kilómetros recorridos por alumnos en la universidad. Determinar la media.

# CLASE	CLASES		fi	fa	Marca de clase MC	MC* fi
	Li	Ls				
1	0.1	18	15	15	9.05	135.75
2	18.1	36	14	29	27.05	378.7
3	36.1	54	28	57	45.05	1261.4
4	54.1	72	26	83	63.05	1639.3
5	72.1	90	17	100	81.05	1377.85
			100			4793.00
<b>MEDIA</b>						<b>47.93</b>

### Mediana

#### Fórmula

Cuando se quiere calcular la mediana en datos agrupados por intervalos se tiene que buscar el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas, es decir, es necesario localizar el intervalo donde se encuentre  $\frac{N}{2}$

para lo cual se utiliza la siguiente fórmula:

$$Me = Li + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$$

$L_i$  = Límite inferior del renglón en donde debe estar la mediana

$F_{i-1}$  = Frecuencia acumulada anterior al renglón de la mediana

$f_i$  = Frecuencia del renglón de la mediana

$a_i$  = Tamaño del intervalo

### Ejemplo

Pasos para buscar la mediana

Recuerda que la mediana representa el valor que divide a los datos en la mitad exacta, es decir, a la derecha del valor de la mediana se encuentran el 50% de los datos y a la izquierda de dicho valor el otro 50%, por lo que para una distribución con datos agrupados se deben seguir los siguientes pasos:

**1** Ubicar la clase de la mediana, para ello se debe buscar en qué clase se encuentra  $\frac{N}{2} = \frac{100}{2} = 20$

**2** Ubicar en la frecuencia acumulada el dato 20

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

- 3** Ubicar el límite inferior de la clase de la mediana que es igual a 6.63
- 4** Ubicar la frecuencia de la clase de la mediana que es igual a 12
- 5** Ubicar la frecuencia acumulada anterior a la clase de la mediana es igual a 14.
- 6** Ubicar la amplitud de la clase que es 21,

El siguiente esquema representa algunos de los pasos descritos:

Li	Ls	Fi	Fa
5.97	6.18	2	2
6.19	6.4	5	7
6.41	6.62	7	14
<b>6.63</b>	6.84	12	26
6.85	7.06	8	34
7.07	7.28	6	40

Diagram annotations: A blue arrow labeled '3' points to the Li value 6.63. A yellow arrow labeled '5' points to the Fa value 14. A red arrow labeled '2' points to the Fi value 12. A yellow arrow labeled '4' points to the Fa value 34.

Sustituyendo en la fórmula de la mediana se tiene que el valor de la mediana es 6.73

### Moda

#### Fórmula

La moda es el valor del dato que más veces se repite, esto es, el valor cuya frecuencia absoluta es mayor, y se denota como **Mo**. Algunas veces el valor que más se repite puede no ser único, es decir, puede haber dos o más datos que aparezcan con la misma frecuencia absoluta, siendo ésta la mayor. En esas ocasiones se habla de poblaciones o muestras bimodales cuando existen dos modas o multimodales si existen más de dos.

Por ejemplo, si se toma una muestra de hombres y mujeres y se miden sus estaturas, se tienen dos modas.

Cuando la distribución de datos es por intervalos de clase, primero se localiza el intervalo que tiene mayor frecuencia absoluta y se utiliza la siguiente fórmula para calcular la moda:

$L_i = \text{Límite inferior del renglón en donde debe estar la moda}$
--

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

$Mo = Li + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot a_i$	$f_i$ = Frecuencia del renglón de la moda
	$f_{i+1}$ = Frecuencia ulterior al renglón de la moda
	$f_{i-1}$ = Frecuencia anterior al renglón de la moda
	$a_i$ = Tamaño del intervalo

### Ejemplo

Como se mencionó con anterioridad, la moda corresponde al valor o valores, si es multimodal, que más se repiten en una distribución; para el caso de datos agrupados se deben seguir los siguientes pasos para obtener el valor de la moda.

Pasos para buscar la moda

- 1 Ubicar la clase de la moda y ésta es la clase donde se tienen más datos, es decir, hay 12 datos entre 6.63 y 6.84, como puedes observar en la cuarta fila
- 2 Ubicar el límite inferior de la clase de la moda, el cual es 6.63
- 3 Calcular  $(f_i - f_{i-1}) = 12 - 7 = 5$
- 4 Calcular  $(f_i - f_{i+1}) = 12 - 8 = 4$
- 5 Ubicar la amplitud de la clase 21

El siguiente esquema representa algunos de los pasos descritos:

Li	Ls	Fi	Fa
5.97	6.18	2	2
6.19	6.4	5	7
6.41	6.62	7	14
6.63	6.84	12	26
6.85	7.06	8	34
7.07	7.28	6	40

Sustituyendo en la fórmula de la moda se tiene 6.74

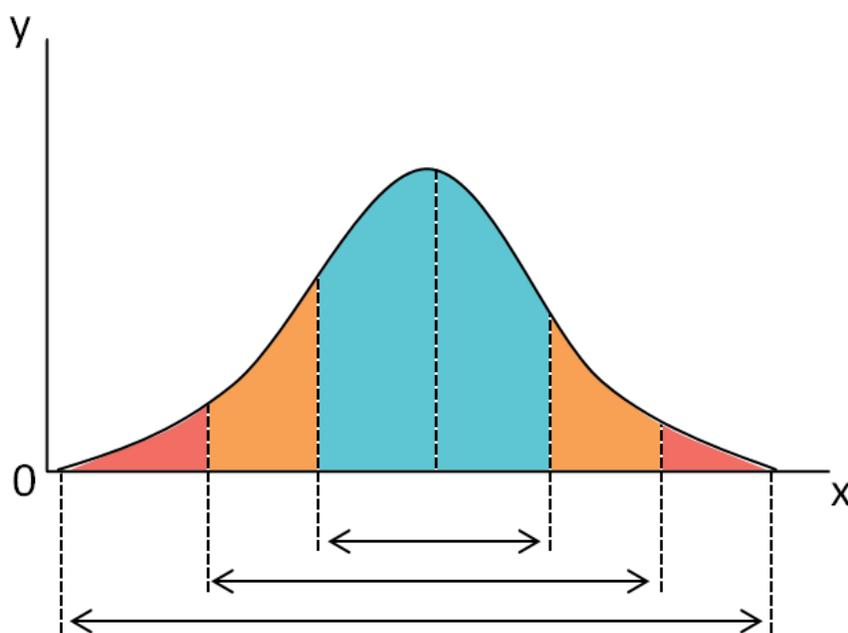
# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### 3.3. Medidas de dispersión

A diferencia de las medidas de tendencia central que miden acumulaciones mediante un solo punto, las medidas de dispersión miden el grado de separación o alejamiento que tiene una variable estadística en torno a una medida de posición o tendencia central. Dicho grado de separación indica lo representativa que es la medida de posición con respecto al conjunto total de datos. A mayor dispersión menor representatividad de la medida de posición y viceversa.

Las medidas de dispersión más comunes son: el recorrido, la varianza y la desviación estándar.



#### 3.3.1. Datos no agrupados (varianza y desviación estándar)

Al igual que las medidas de tendencia central, las medidas de dispersión se pueden obtener a partir de datos agrupados o no agrupados y de manera análoga para datos poblacionales o bien muestrales como a continuación se mostrará.

##### Varianza

La varianza mide la mayor o menor dispersión de los valores de la variable respecto a la media aritmética. Siempre es mayor o igual que cero y menor que infinito. Se define como la media de los cuadrados de las diferencias del valor de los datos menos la media aritmética de éstos.

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

Las fórmulas de la varianza para datos no agrupados son:

En una población	En una muestra
$S^2 = \frac{\sum_{i=1}^n (x_i - m)^2}{N}$	$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$

Para obtener la varianza se realiza la sumatoria de cada valor menos la media y se eleva al cuadrado y el resultado se divide ya sea entre el valor poblacional ( $N$ ), o bien el muestral menos 1, que corresponde a:  $n-1$ .

### Desviación típica o estándar

La desviación típica muestra qué tan alejado está un dato del valor de la media aritmética, es decir, la diferencia que hay entre un dato y la media aritmética. Se denota como  $S$  o  $\sigma$ , según se calcule en una muestra o en toda la población, respectivamente. Se define como la raíz cuadrada positiva de la varianza.

Las fórmulas de la desviación típica o estándar para datos no agrupados son:

En una población	En una muestra
$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$

Es decir que, al valor de la varianza, ya sea poblacional o muestral, se le aplica la raíz cuadrada y se obtiene la desviación típica o estándar.

### 3.3.2. Datos agrupados (varianza y desviación estándar)

#### Varianza para datos agrupados por intervalos

Las fórmulas para calcular la varianza en datos agrupados por intervalos son las siguientes:

En una población	En una muestra
------------------	----------------

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

$$\sigma^2 = \frac{\sum_{i=1}^N f_i (Mc_i - \mu)^2}{N} \quad S^2 = \frac{\sum_{i=1}^n f_i (Mc_i - \bar{x})^2}{n}$$

En este caso se realiza la sumatoria de cada marca de clase menos la media (ya sea poblacional o muestral, según sea el caso) y se eleva al cuadrado, al final se divide entre la población o bien la muestra, según se trate.

### Desviación típica o estándar en datos agrupados por intervalos

Las fórmulas para calcular la desviación típica o estándar en datos agrupados por intervalos son las siguientes:

En una población	En una muestra
$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N f_i (Mc_i - \mu)^2}{N}}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n f_i (Mc_i - \bar{x})^2}{n - 1}}$

De manera análoga al resultado de la varianza se le aplica la raíz cuadrada y se obtiene la desviación estándar, ya sea para una población o bien una muestra.

### Ejemplo

Los siguientes datos se refieren al diámetro en pulgadas de un engrane.

6	6.25	6.5	6.65	6.75	6.75	7	7.1
6	6.25	6.5	6.7	6.75	7	7	7.15
6.25	6.5	6.5	6.7	6.75	7	7	7.15
6.25	6.5	6.5	6.75	6.75	7	7	7.25
6.25	6.5	6.65	6.75	6.75	7	7.1	7.25

Obtenga la media, la varianza y la desviación estándar.

LI	LS	FI	MC	FA	FR	FRA	F. Porcentual	(360°F I)/N	FI*MC	(MC- MEDIA) ^2*FI
5.97	6.18	2	6.08	2	0.05	0.05	5	18	12.15	0.83
6.19	6.4	5	6.30	7	0.13	0.18	12.5	45	31.475	0.90
6.41	6.62	7	6.52	14	0.18	0.35	17.5	63	45.605	0.29
6.63	6.84	12	6.74	26	0.30	0.65	30	108	80.82	0.00
6.85	7.06	8	6.96	34	0.20	0.85	20	72	55.64	0.45
7.07	7.28	6	7.18	40	0.15	1.00	15	54	43.05	1.25

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

N	40					100	360	268.74	3.72
---	----	--	--	--	--	-----	-----	--------	------

Media = 6.72  
Varianza = 0.093  
Desviación estándar = 0.305

### Cierre de la Unidad



Indicador

Fuente: Master isolated images, 2011.  
[freedigitalphotos.net](http://freedigitalphotos.net)

Al terminar esta unidad has podido tener una visión general de la importancia que tienen las medidas de tendencia central y de dispersión para comprender el comportamiento de los datos y poder ubicar qué tan cercano o alejado se encuentra un valor cualquiera de las medidas de tendencia central, así como la variabilidad y dispersión de los datos en conjunto de una distribución, de manera que puedas interpretar mejor tu problema.

### Fuentes de consulta

#### Básica

- Ibarra, O. M. (2006). *Estadística para la Administración Turística*. México: Trillas.
- Levin, R. y Rubin, D. (2010). *Estadística para administración y economía*. México: Pearson.
- Montgomery, D. C. y Runger, G. C. (1996). *Probabilidad y Estadística aplicadas a la ingeniería*. (Cuarta edición). México: McGraw-Hill.
- Shao, S. P. (2007). *Estadística para Economistas y Administradores de Empresas*. México: Herrero Hermanos.
- Walpole, R. E., Myers, R. H. et al. (2007). *Probabilidad y Estadística para Ingeniería y ciencias*. (Octava edición). México: Pearson Educación.

# Estadística Básica

## Unidad 3. Muestreo, medidas de tendencia central y de dispersión

### Fuentes electrónicas

- Asesor Juan Manuel. Tutorial Excel-Función Aleatorio- Random.:  
<http://www.youtube.com/watch?v=kr56Sc4YygY>
- Espinoza Salvadó, I. (2016). Tipos de muestreo.  
<http://www.bvs.hn/Honduras/Embarazo/Tipos.de.Muestreo.Marzo.2016.pdf>
- Matemóvil. (2019). Población y muestra - ejemplos y ejercicios resueltos (Video). Youtube.  
<https://www.youtube.com/watch?v=gl9EEbT7viM>
- Ruiz, M. Mediana para datos agrupados.  
<https://www.youtube.com/watch?v=02Q70BiSQgo>
- Ruiz, M. Moda para Datos Agrupados.  
<https://www.youtube.com/watch?v=Y1Ja2C8VtAE>
- Salazar, P. C. & Del Castillo, G. S. (2018). Fundamentos básicos de estadística.  
<http://librodigital.sangregorio.edu.ec/librosusgp/B0009.pdf>
- Vitutor (2012). *Estadística y probabilidad*.  
<http://www.vitutor.com/estadistica.html>