



GUIA UNIDAD: ESTADÍSTICA BÁSICA

La Estadística consiste en un conjunto de técnicas y procedimientos que permiten recoger datos, presentarlos, ordenarlos y analizarlos, de manera que a partir de ellos se puedan inferir conclusiones.

Conceptos básicos

- 1) La **población** es un conjunto de objetos o de individuos que se desea estudiar y que, a su vez, presentan una característica que interesa medir.

Ejemplo: Estudiantes del IN



- 2) Se llama **muestra** a un subconjunto representativo de la población que se desea estudiar.

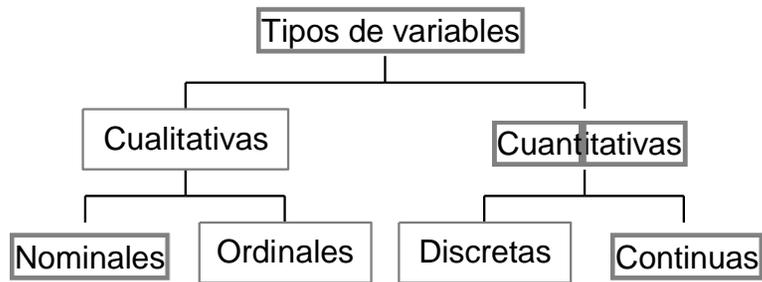
Ejemplo: estudiantes de los octavos básicos del IN



- 3) **Muestra aleatoria:**

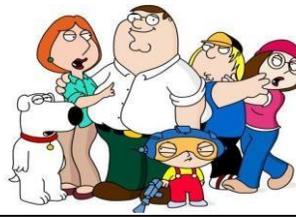
Es una muestra al azar. Para que se considere propia y representativa de la población, deberá ser al azar.

- 4) Una **variable** estadística corresponde a la o las características que se miden en la muestra. Su expresión numérica es el **dato**.



Variables cuantitativas tienen valores numéricos que representan medidas (largo, peso, etc.) o frecuencias (número de). Tiene sentido realizar operaciones numéricas con estas variables. Además distinguimos dentro de las variables cuantitativas las discretas y las continuas.

Una variable **discreta** solo se mide por medio de números enteros.
Ejemplos: número de páginas de un libro, número de hijos, etc.

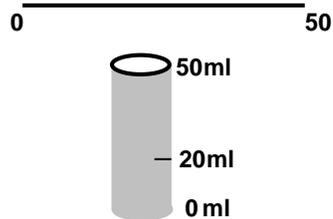


Una variable **continua** puede tomar cualquier valor de un número real en un intervalo dado.

Ejemplos:

Peso (masa) de un individuo.

Cantidad de agua en un vaso de 50 ml.



Variables cualitativas son aquellas que no se pueden medir numéricamente, están relacionadas con características. A sus valores se les llama categorías. Se clasifican en:

Nominal: son aquellas en las cuales no existe ninguna ordenación. Ejemplos: Marca de auto, Sexo, Religión, estado civil, etc.

Ordinal: tiene asociado un orden.

Ejemplo: Nivel educacional, Estado nutricional, Nivel Socioeconómico, etc.

5) **Dato:** es el valor de la variable para cada elemento perteneciente a la población o a la muestra.

ACTIVIDAD 1:

Determine qué tipo son las siguientes variables. Si son variables cualitativas (nominal u ordinal) o cuantitativas (discretas o continuas).

- a) Marca de automóvil
- b) Duración de un compacto (segundos)
- c) Número de temas de un compacto
- d) Nivel educacional (básica, media, universitaria)
- e) Temperatura al mediodía en Talca (grados Celcius)
- f) Estado civil (soltero, casado, divorciado, viudo)
- g) Cantidad de lluvia en un año en Talca (mm^3)
- h) La nacionalidad de una persona.
- i) Número de litros de agua contenidos en un depósito.
- j) Número de libros en un estante de librería.
- k) Suma de puntos tenidos en el lanzamiento de un par de dados.
- l) La profesión de una persona.
- m) El área de las distintas baldosas de un edificio.

ORDENACIÓN DE LOS DATOS

Al ordenar los datos correspondientes a un cierto estudio, es usual agruparlos en clases o categorías, para lo cual se utilizan tablas de distribución de frecuencias.

Distribución de frecuencias

Frecuencia absoluta (f): es el número de veces que aparece o se repite un cierto valor en la variable de medición.

Frecuencia absoluta acumulada (F): representa el número de datos cuyo valor es menor o igual al valor considerado. Se obtiene sumando sucesivamente las frecuencias absolutas.

Frecuencia relativa (f_r): representa la razón de ocurrencia respecto al total. Se calcula como el cociente entre la frecuencia absoluta y el tamaño total de la muestra. La suma de todas las frecuencias relativas da como resultado 1.

Frecuencia relativa porcentual ($f_r\%$): corresponde a la frecuencia relativa expresada en porcentaje. Se calcula como el producto de la frecuencia relativa por 100.

Frecuencia relativa porcentual acumulada ($F_r\%$): corresponde a la frecuencia relativa acumulada expresada como porcentaje. Se calcula el producto de la frecuencia acumulada por 100.

Tablas de distribución de frecuencias para datos no agrupados

Una tabla de distribución de frecuencias es una representación de la información obtenida de una muestra o población, en relación con los valores que puede tomar una variable.

Ejemplo: las siguientes son las notas obtenidas en una prueba por los 30 alumnos de un curso

6-5-5-4-6-3-2-2-7-6-1-5-6-6-6
7-2-3-4-4-5-4-5-5-6-6-6-7-7-5

Con la ayuda de tu profesor(a) completa la tabla de frecuencias que corresponde a la situación.

Notas	f _i	F _i	f _r	f _r %	F _r %
1					

2					
3					
4					
5					
6					
7					

De acuerdo a la información de la tabla, responde:

- ¿cuántos estudiantes tuvieron nota inferior a 4?
- ¿Cuántos estudiantes tuvieron a lo menos una nota 4?
- ¿Qué porcentaje de los estudiantes tuvo nota entre 4 y 7?
- ¿Qué representa f_6 ?
- ¿Qué porcentaje de estudiantes reprobó la prueba?

ACTIVIDAD 2:

1) En la tabla se muestran las edades de 25 jugadores de vóleybol.

Edad (años)	f_i	F_i	f_r	$f_r\%$	$F_r\%$
18	3				
19	5				
20	7				
21	5				
22	2				
23	2				
24	1				

Responde:

- Completa la tabla de frecuencias.
 - ¿Cuántos jugadores tienen 22 años o menos?
 - ¿Qué porcentaje de jugadores tienen 19 años?
 - ¿Qué porcentaje de jugadores tienen más de 20 años y menos de 23 años?
 - ¿hay más jugadores que tienen menos de 21 años o que tienen 21 años o más?
- 2) En una empresa se requiere recopilar información acerca de la cantidad de teléfonos móviles que hay en un grupo de hogares. Para ello, se realizó una encuesta a 40 empleados, de distintas familias, obteniéndose la siguiente información:

4 - 3 -1 -2 - 3 -4 -0 - 5- 1- 3
1 - 0 -4 -5 - 4 -1 -2 - 1- 3- 3
3 - 4 -3 -3 - 2 -5 -2 - 3- 3- 2
2 - 1 -2 -3 - 2 -2 -3 - 1- 4- 2

- ¿Cuál es la población de estudio?
- ¿Cuál es la muestra de estudio?
- ¿Cuál es la variable de estudio?, ¿de qué tipo es esta variable?

- d) Construye una tabla de frecuencias con los datos.
- e) ¿Qué porcentaje de los hogares tienen 2 teléfonos móviles?
- f) ¿en cuántos hogares hay 3 o menos teléfonos móviles?
- g) ¿Qué porcentaje de los hogares no tienen teléfono móvil?
- h) ¿en cuántos hogares hay 4 o 5 teléfonos móviles?

Tablas de frecuencias para datos agrupados

Si los datos que se recolectan para estudiar una variable cuantitativa son numerosos o se consideran muchos valores distintos, se pueden agrupar en intervalos.

Intervalo: un intervalo $[a, b[$ es el conjunto de los números mayores o iguales que **a** y menores que **b**, donde **a** es el *límite inferior* y **b**, el *límite superior* del intervalo.

Amplitud de un intervalo: es la diferencia entre **b** y **a**.

Rango: diferencia entre el mayor y el menor valor de una variable.

$$\text{Rango} = X_{\text{máx}} - X_{\text{mín}}$$

Marca de clase (M_c): es un valor representativo de cada intervalo. Este valor corresponde al punto medio del intervalo. Se calcula como la suma del límite inferior (**a**) y el límite superior (**b**) del intervalo, dividido en dos.

$$M_c = \frac{a + b}{2}$$

Construcción de una tabla de frecuencias para datos agrupados

Ejemplo: Construir una tabla de Distribución de Frecuencia en intervalos de clase con las estaturas, en centímetros, de 80 alumnos de un colegio.

150	152	153	153	154	155	156	157	150
153	169	154	155	156	157	150	152	153
154	155	158	157	152	153	153	154	155
156	157	152	153	153	154	155	155	154
158	161	161	162	165	162	168	170	158
162	162	165	166	168	172	159	161	174
163	165	166	168	172	159	161	167	163
167	169	172	160	161	162	163	166	162
154	157	156	161	166	155	152	165	

1° Paso: Determinar Rango:

$$R = \text{Mayor Valor} - \text{Menor Valor}$$

$$R = 174 - 150 = 24$$

2° Paso: Determinar el Número de Intervalos:

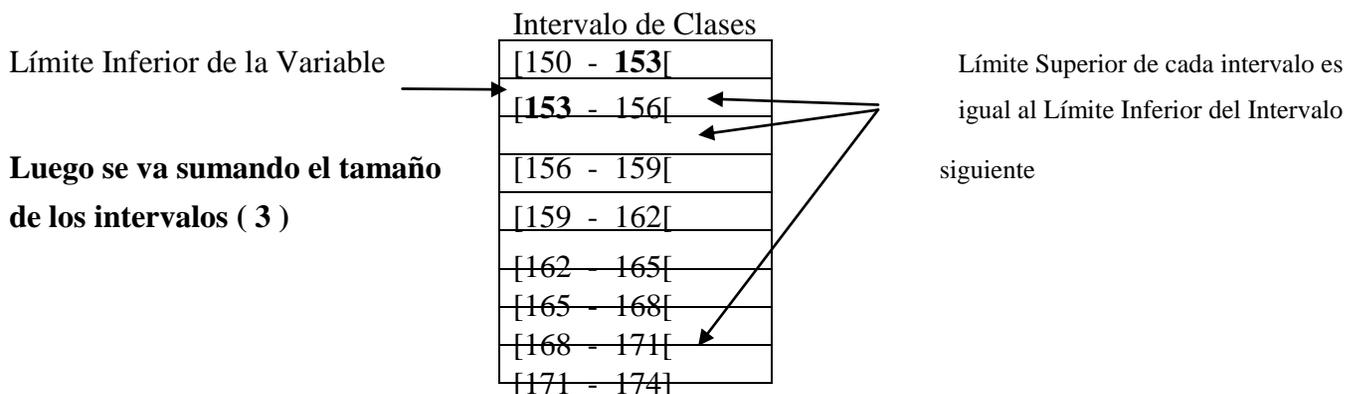
Consideremos el Valor 8 (Además es Factor de 24)

3° Paso: Determinemos el Tamaño de los Intervalos, amplitud (c):

$$c \approx \frac{R}{N \text{ de Intervalos}} = \frac{24}{8} \approx 3$$

Nota: Si el tamaño del intervalo no resulta ser un número entero, **a veces** (si la variable es cuantitativa discreta) conviene aproximarlos al entero que corresponda.

4° Paso: Fijar el Límite Inferior y el Límite Superior de cada Intervalo.



un dato igual al límite superior de un intervalo corresponderá al intervalo siguiente, por ejemplo: el valor de 153 corresponde al intervalo [153 - 156[y no al intervalo [150 - 153[

5° Paso: Determinar la Marca de Clase y completar las frecuencias.

Intervalo de Clases	Mc	fi	Fi	fr	fr%	Fr%
[150 - 153[151,5	8	8		10%	
[153 - 156[154,5		30			37,5%
[156 - 159[157,5	12				
[159 - 162[160,5				11,25%	63,75%
[162 - 165[163,5	9				
[165 - 168[166,5		70		12,5%	
[168 - 171[169,5					
[171 - 174]	172,5		80			
Totales →		n = 80			100%	

ACTIVIDAD 3:

1) Las edades de 195 alumnos de un colegio son las siguientes:

13	15	19	14	15	17	17	18	16	14	17	14	19
14	15	15	13	16	17	18	19	17	13	17	13	20
16	17	16	15	16	18	17	20	16	14	18	14	17
18	16	16	13	14	15	16	18	17	15	19	16	17
19	18	14	14	13	16	15	17	15	16	15	18	16

14	15	13	18	16	15	14	15	17	16	15	20	17
16	17	13	18	17	13	15	14	18	17	14	14	16
17	16	15	15	18	14	17	15	19	17	13	15	18
18	14	17	16	19	16	16	13	18	18	18	16	17
19	13	17	14	17	13	17	14	15	15	19	17	17
14	15	15	17	14	15	18	13	16	15	16	19	17
15	16	17	16	16	17	19	16	15	13	17	17	16
16	17	14	15	17	15	15	18	17	14	16	15	13
18	16	18	15	18	16	20	17	18	16	16	14	18
19	14	18	13	16	17	14	15	19	17	18	16	15

- En tu cuaderno construye una tabla de distribución de frecuencias y contesta las siguientes preguntas (aproxima a tres decimales)
- ¿Cuántos alumnos tienen 17 años?
- ¿Cuántos alumnos tienen 16 años o menos de 16?
- ¿Cuántos alumnos tienen menos de 15 años?
- ¿Qué % de alumnos tiene 14 años?
- ¿Qué % de alumnos tiene 20 años?
- ¿Qué % de alumnos tienen menos de 19 años?
- ¿Qué % de alumnos tienen 17 años o menos?

2) Construye en tu cuaderno la tabla de distribución de frecuencias con los datos correspondientes a la duración en horas de 80 ampollitas fabricadas por una industria.

295	305	308	291	298	323	328	300	297	297	294	295	295	294	299	290
322	304	304	299	324	299	320	320	302	320	300	324	300	301	311	304
323	321	322	313	310	317	311	318	301	310	308	303	312	306	314	307
313	315	312	314	318	311	310	310	314	317	317	306	307	318	319	317
319	319	305	309	306	314	309	310	319	315	312	308	313	316	311	319

De acuerdo con la tabla anterior, contesta las siguientes preguntas:

- ¿Cuántas ampollitas tuvieron una duración entre 300 y 305 horas?
- ¿Cuántas tuvieron una duración 305 o más horas?
- ¿Cuántas tuvieron una duración menor que 320 horas?
- ¿Qué porcentaje representan las ampollitas que duraron entre 310 y 315 horas?
- ¿Qué porcentaje representan las ampollitas que duraron entre 300 y 325 horas?
- ¿Qué porcentaje representan las ampollitas que duraron 310 o menos horas?
- ¿Qué porcentaje representan las ampollitas que duraron menos de 325 horas?

TEST

Marca la opción correcta en cada caso:

1) La tabla de frecuencias muestra el tiempo de reacción, en minutos, de 40 personas luego de aplicarles un medicamento. ¿Cuál de las siguientes afirmaciones son FALSAS?

Tiempo de reacción al medicamento			
X	f	F	f%
[10 - 15[8	8	20
[15 - 20[9	17	22,5
[20 - 25[12	29	30
[25 - 30[11	40	27,5

- I. La marca de clase del tercer intervalo es 25.
 II. El 30% de las personas reacciona entre los 20 y 30 minutos.
 III. El 42,5% de las personas reacciona antes de los 20 minutos.

- A) Sólo I
 B) Sólo II
 C) Sólo III
 D) Sólo I y II
 E) I, II y III

Utilizando la información de la tabla responde las preguntas 2, 3 y 4.

Notas de 1° medio		
	f	F
[6,0; 7,0]	14	14
[5,0; 6,0[15	29
[4,0; 5,0[8	37
[3,0; 4,0[2	39
[2,0; 3,0[1	40
[1,0; 2,0[0	40

2) ¿Cuántos estudiantes obtuvieron al menos una nota 4,0?

- A) 14
 B) 29
 C) 37
 D) 39
 E) 40

3) ¿Cuántos estudiantes tiene el curso?

- A) 35
 B) 37
 C) 39
 D) 40
 E) Ninguna de las anteriores.

4) ¿Qué porcentaje de los estudiantes tiene una nota menor que 4,0?

- A) 3%
 B) 5%
 C) 7,5%
 D) 8,3%
 E) 9,6%

En la siguiente tabla se muestran las horas de entrenamiento diario de un grupo de futbolistas profesionales:

Horas	Cantidad de
7	futbolistas 150

8	230
9	225
10	100

5) ¿Cuántas horas como máximo entrena el 40% de ellos?

- A) 7 horas
- B) 8 horas
- C) 9 horas
- D) 10 horas
- E) Más de 10 horas

La siguiente información muestra los puntajes obtenidos por 50 alumnos de IV° medio en la PSU de matemáticas

320	450	480	500	510	540	570	610	670	700
350	450	480	500	520	540	580	620	670	700
400	460	480	510	520	540	580	630	680	710
400	460	490	510	520	550	610	630	690	720
430	470	500	510	530	570	610	640	690	750

6) Si se ordenan los datos en una tabla de frecuencias con 10 intervalos, ¿cuál es la amplitud de cada intervalo?

- A) 40
- B) 41
- C) 42
- D) 43
- E) 44

Las siguientes son las distancias que 30 estudiantes elegidos al azar de una escuela recorren diariamente de ida y vuelta para ir a la escuela (en cuadras).

12	30	10	11	18	26	34	18	8	12
26	14	5	22	4	25	9	10	6	21
18	18	9	16	44	23	4	13	36	8

7) Si la amplitud del intervalo es 5, ¿cuántos intervalos tendría la tabla de frecuencias?

- A) 4
- B) 5
- C) 6
- D) 7
- E) 8

8) La tabla adjunta muestra las frecuencias (f) de las notas en la prueba de matemática, obtenidas por los alumnos de 8° Básico de un colegio, ¿Cuáles de las siguientes afirmaciones son verdaderas?

- I) El 75% del curso obtuvo una nota igual o inferior a 5,5
- II) La moda corresponde a la nota 5,0
- III) El 15% del curso obtuvo la nota 4,5
- IV) El 50% del curso obtuvo nota superior a 5.0

- A) Sólo II y III
- B) Sólo III y IV
- C) Sólo I, II y III
- D) Sólo I, II y IV
- E) Sólo II, III y IV

Nota	f
3,0	3
3,5	5
4,0	4
4,5	6
5,0	7
5,5	5
6,0	4
6,5	4
7,0	2
	40