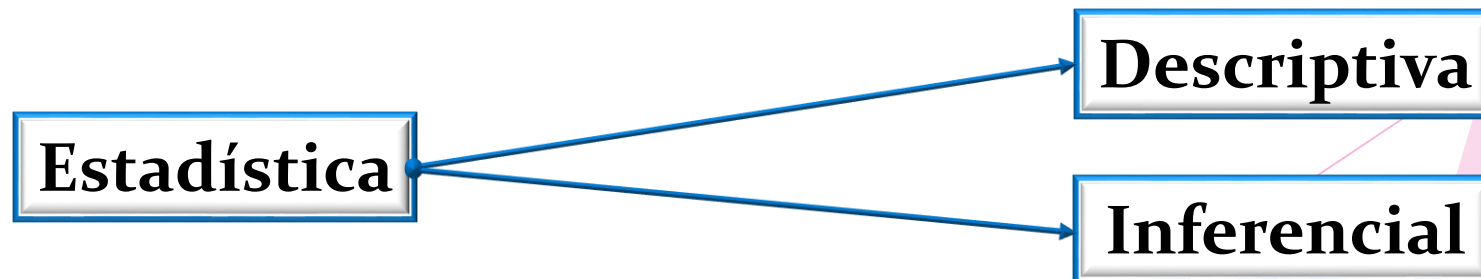


ORGANIZACIÓN DE LA INFORMACIÓN

Penna – Cobos – Vázquez Ferrero – Ulagnero

ESTADÍSTICA: CONCEPTOS Y DEFINICIÓN

- La Estadística es la ciencia que se encarga de recoger, organizar e interpretar datos.
- La Estadística es fundamental para muchas ramas de la ciencia desde la medicina a la economía. Pero, y en lo que a nosotros atañe, es esencial para interpretar los datos que se obtienen de la investigación científica.
- Siendo el objetivo principal de la Estadística obtener conclusiones de la investigación empírica, por eso se hace tan necesaria para todo aquello que requiera analizar datos y diseñar experimentos.
- Podemos, entonces, definir a la Estadística como *“la ciencia de los datos”*.



ESTADÍSTICA DESCRIPTIVA

- Debido a que en estadística no trabajamos con UA aisladas sino con conjuntos de ellas, siempre es necesario resumir la información de forma tal que sea comprensible mediante la lectura y para poder obtener un significado.
- Así, la Estadística Descriptiva nos dará una serie de procedimientos dirigidos a organizar, a resumir y a sintetizar la información para que podamos interpretarla y extraer conclusiones a partir de los datos que obtenemos ya que, de otra manera, serían ininteligibles.

Por ejemplo, si una Escuela presenta una lista con distintos datos de alumnos de dicho Establecimiento Educativo, sería conveniente poder agruparlos, ya sea por género, rango etario, año académico, edad, etc., y organizar esa información en tablas de frecuencias o gráficos.

ESTADÍSTICA INFERENCIAL

- Una vez que disponemos de una síntesis de la información que hemos recogido de un conjunto de UA, nos interesa otro problema: el de preguntarnos si aquello que observamos vale también para otras UA, a las que no hemos observado.

Por ejemplo, una situación cotidiana para los/as estudiantes: supongamos que diferenciamos dos grupos, uno que se exige más a la hora de un examen final y el otro menos. Mediante un test estandarizado nos damos cuenta que los/as estudiantes del primer grupo presentan un mayor nivel de ansiedad a la hora de rendir un examen final. Ahora bien, ¿podemos indicar que la auto-exigencia incide en la ansiedad?, es decir, ¿podemos generalizar nuestro resultado?

- La Estadística Inferencial se ocupará de esto, de decirnos bajo qué condiciones se pueden extender nuestros hallazgos a casos no observados.

ESTADÍSTICA INFERENCIAL (CONT.)

- En todos los casos, cuando sea posible hacer generalizaciones, éstas estarán limitadas a un contexto específico.

Por ejemplo, el análisis que se haga, de la relación entre pobreza y educación en Argentina, puede no ser válido para la población de Brasil. Es decir, se debe explicitar cuál es la población de referencia a la cual es válido extender los resultados que se obtienen.

ALGUNAS CONSIDERACIONES

- Puede considerarse a la descripción como una etapa anterior a la inferencia, ya que esta última no puede lograrse sin una adecuada descripción previa de la información.
- Pero la descripción tiene entidad propia y, como ya hemos visto, en un estudio pueden plantearse objetivos exclusivamente descriptivos, que no se dirijan a generalizar los resultados obtenidos.

ORGANIZACIÓN DE LA INFORMACIÓN

Como se vio en la unidad anterior, el concepto de variable es uno de los más importantes en estadística. Pero, para seguir abordando esta definición, es necesario introducir algunos conceptos (que ampliaremos en unidades posteriores) para completar esta idea.

POBLACIÓN Y MUESTRA

Se denomina **población** (N) al conjunto completo de UA, con alguna característica en común, que es el objeto de nuestro estudio.

Una población puede ser finita o infinita:

- Los habitantes de un país, los planetas del Sistema Solar, la cantidad de alumnos en una facultad, etc., son elementos de una población finita.
- Con respecto a las infinitas: número de gotas de agua, cantidad de granos de arena en la playa, etc.

ORGANIZACIÓN DE LA INFORMACIÓN (CONT.)

Sin embargo, aunque la población sea finita la cantidad de UA que la constituyen puede ser elevada. Para estos casos se hace necesario trabajar sólo con una parte -o un subconjunto- de dicha población. A esta parte o subconjunto la llamamos **muestra (n)**.

- Un ejemplo clásico es cuando tenemos una enfermedad y debemos hacernos un análisis de sangre. ¿Se imaginan si tuvieran que sacarnos toda la sangre para poder analizarla?

TABLAS DE FRECUENCIAS

Cuando se define una variable deben indicarse el conjunto de categorías que le corresponden, aunque a veces esto está implícito.

- Por ejemplo, si se trata del mayor nivel de escolaridad alcanzado, pueden considerarse las siguientes categorías: ninguno, primario, secundario y terciario.

PROPIEDADES DE LAS CATEGORÍAS

- **Exclusión mutua:** Cada categoría excluye a todas las demás. Dicho de otra manera, si a una UA le corresponde una categoría, entonces sabemos que no le corresponde ninguna otra.
- **Exhaustividad:** Este requisito que solicitaremos a las categorías de una variable es que agoten todas las posibilidades de variación, es decir, que cada UA tenga una categoría de pertenencia.

TABLAS DE FRECUENCIAS (CONT.)

Es importante tener en cuenta que una Tabla de Frecuencia (como así también un Gráfico) debe poseer, para su correcta interpretación:



Título: debe describir claramente lo que la Tabla de Frecuencias (o Gráfico) ilustra.

Referencias: se tienen que aclarar las posibles abreviaturas que sean utilizadas.

Fuente: si los datos no proviene de una fuente primaria, se debe consignar de dónde fueron obtenidos.

NIVEL DE MEDICIÓN NOMINAL

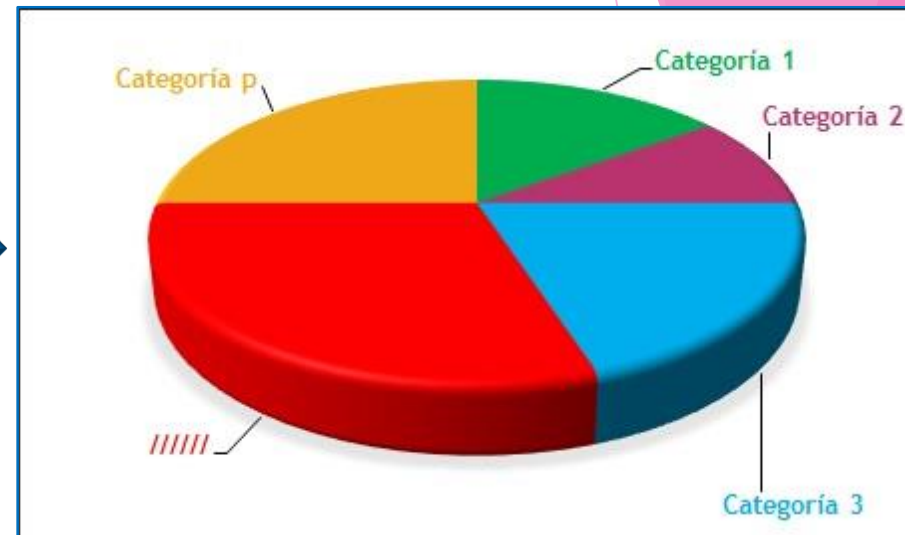
Título

SC	f	f'	f'%
Categoría 1	f_1	$f'_1 = \frac{f_1}{n}$	$f'_1\% = \frac{f_1}{n} \times 100$
Categoría 2	f_2	$f'_2 = \frac{f_2}{n}$	$f'_2\% = \frac{f_2}{n} \times 100$
Categoría 3	f_3	$f'_3 = \frac{f_3}{n}$	$f'_3\% = \frac{f_3}{n} \times 100$
//////
Categoría p	f_p	$f'_p = \frac{f_p}{n}$	$f'_p\% = \frac{f_p}{n} \times 100$
	n	1	100

Gráfico circular
o de sectores.



Título



Referencias

Fuente

Referencias

Fuente

Siendo:

- SC = sistema de clasificación. Variable estudiada.
- f = frecuencia absoluta. Número de UA por cada categoría.
- f' = frecuencia relativa. Peso relativo de cada categoría respecto al total.
- f'% = frecuencia relativa porcentual. Porcentaje de cada categoría.

NIVEL DE MEDICIÓN ORDINAL

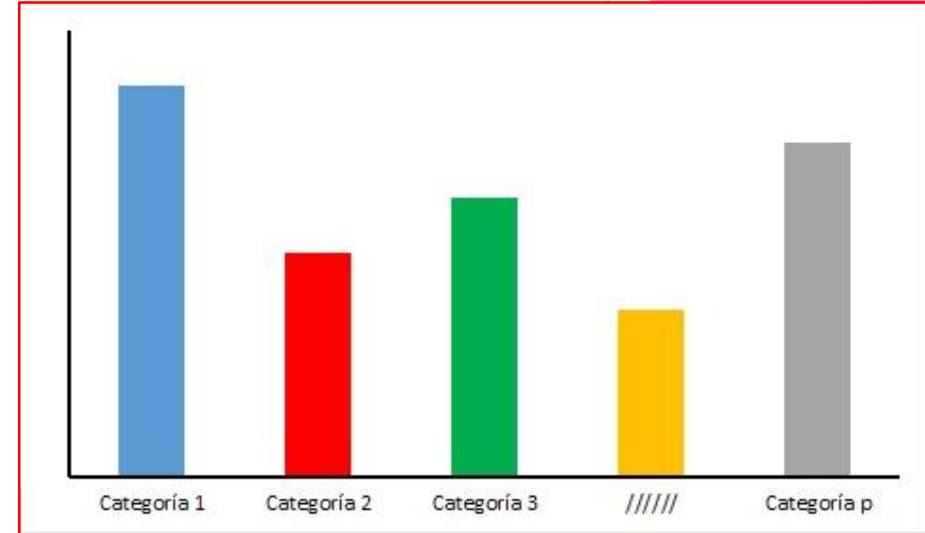
Título

SC	f	f'	f'%
Categoría 1	f_1	$f'_1 = \frac{f_1}{n}$	$f'_1\% = \frac{f_1}{n} \times 100$
Categoría 2	f_2	$f'_2 = \frac{f_2}{n}$	$f'_2\% = \frac{f_2}{n} \times 100$
Categoría 3	f_3	$f'_3 = \frac{f_3}{n}$	$f'_3\% = \frac{f_3}{n} \times 100$
//////
Categoría p	f_p	$f'_p = \frac{f_p}{n}$	$f'_p\% = \frac{f_p}{n} \times 100$
	n	1	100

Gráfico de barras



Título



Referencias

Fuente

Referencias

Fuente

Siendo:

- SC = sistema de clasificación. Variable estudiada.
- f = frecuencia absoluta. Número de UA por cada categoría.
- f' = frecuencia relativa. Peso relativo de cada categoría respecto al total.
- f'% = frecuencia relativa porcentual. Porcentaje de cada categoría.

VARIABLE DISCRETA

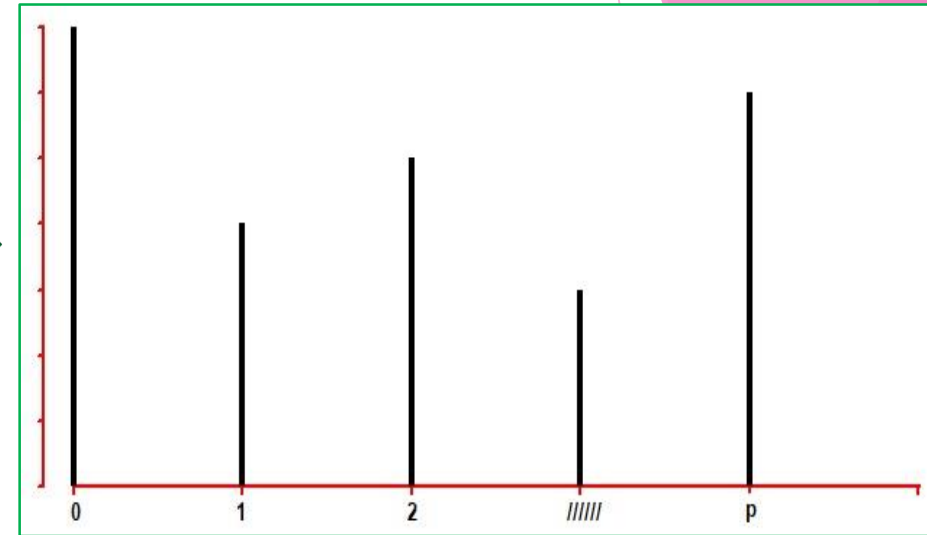
Título

SC	f	f'	f'%
0	f_1	$f'_1 = \frac{f_1}{n}$	$f'_1\% = \frac{f_1}{n} \times 100$
1	f_2	$f'_2 = \frac{f_2}{n}$	$f'_2\% = \frac{f_2}{n} \times 100$
2	f_3	$f'_3 = \frac{f_3}{n}$	$f'_3\% = \frac{f_3}{n} \times 100$
////
p	f_p	$f'_p = \frac{f_p}{n}$	$f'_p\% = \frac{f_p}{n} \times 100$
	n	1	100

Gráfico de líneas



Título



Referencias

Fuente

Siendo:

- SC = sistema de clasificación. Variable estudiada.
- f = frecuencia absoluta. Número de UA por cada categoría.
- f' = frecuencia relativa. Peso relativo de cada categoría respecto al total.
- f'% = frecuencia relativa porcentual. Porcentaje de cada categoría.

Referencias

Fuente

VARIABLE CONTINUA

Para una variable continua, y debido a que los valores obtenidos de las UA pueden ser todos diferentes, se utilizarán intervalos de clases en el Sistema de Clasificación.

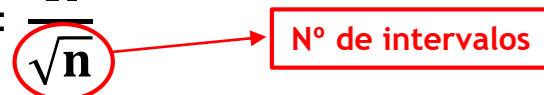
Los pasos a seguir para su construcción, son los siguientes:

- El primer paso es establecer el valor máximo (x_{Max}) y el valor mínimo (x_{Min}) de la distribución, y con ellos determinamos el Rango:

$$R = x_{\text{Max}} - x_{\text{Min}}$$

- El segundo paso es calcular la amplitud de cada uno de los intervalos:

$$i = \frac{R}{\sqrt{n}}$$



Siendo n el tamaño de la muestra.

Luego la Tabla de Frecuencias se construye de la siguiente manera:

VARIABLE CONTINUA (CONT.)

Título

Intervalos ordenados

SC	x'	f	f'	$f'\%$
$[X_{\text{Min}} ; X_{\text{Min}+i})$	$x'_1 = \frac{X_{\text{Min}} + (X_{\text{Min}} + i)}{2}$	f_1	$f'_1 = \frac{f_1}{n}$	$f'_1\% = \frac{f_1}{n} \times 100$
$[X_{\text{Min}+i} ; X_{\text{Min}+2i})$	$x'_2 = \frac{(X_{\text{Min}} + i) + (X_{\text{Min}} + 2i)}{2}$	f_2	$f'_2 = \frac{f_2}{n}$	$f'_2\% = \frac{f_2}{n} \times 100$
$[X_{\text{Min}+2i} ; X_{\text{Min}+3i})$	$x'_3 = \frac{(X_{\text{Min}} + 2i) + (X_{\text{Min}} + 3i)}{2}$	f_3	$f'_3 = \frac{f_3}{n}$	$f'_3\% = \frac{f_3}{n} \times 100$
//////
$[X_{\text{Min}} + p_i ; X_s]$	$x'_p = \frac{(X_{\text{Min}} + p_i) + X_s}{2}$	f_p	$f'_p = \frac{f_p}{n}$	$f'_p\% = \frac{f_p}{n} \times 100$
		n	1	100

$X_s \geq X_{\text{Max}}$

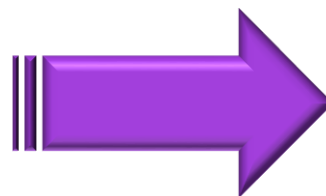
Referencias

Fuente

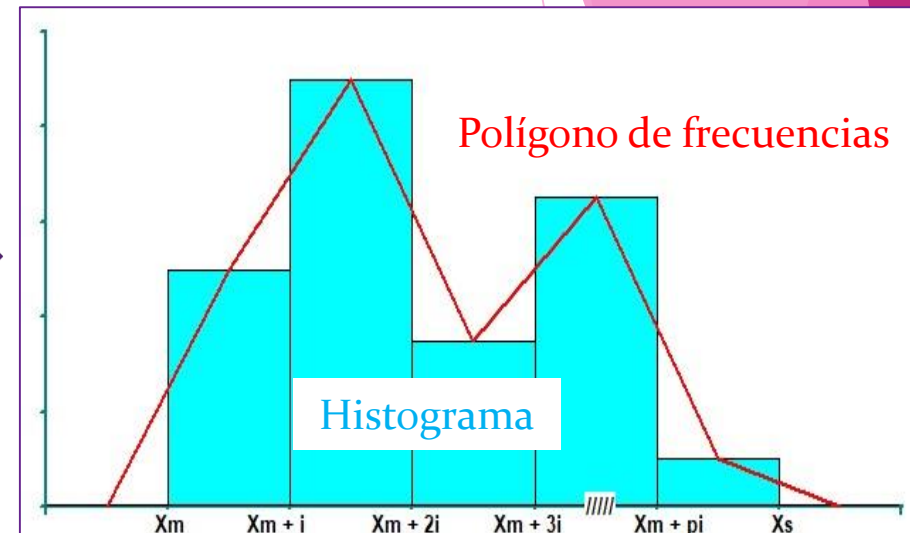
Siendo:

- SC = sistema de clasificación. Variable estudiada.
- x' = punto medio de cada intervalo de clase (IC).
- f = frecuencia absoluta. Número de UA por cada IC.
- f' = frecuencia relativa. Peso relativo de cada IC respecto al total.
- $f'\%$ = frecuencia relativa porcentual. Porcentaje de cada IC.

Histograma /
Polígono de
frecuencias



Título



Referencias

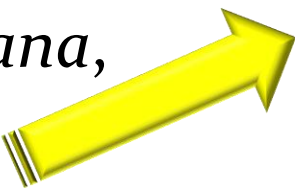
Fuente

EJEMPLO

Un estudiante universitario está interesado en describir las características de estudiantes de una Escuela secundaria XX, de la Ciudad de San Luis.

Para ello toma una muestra aleatoria del 10% de estudiantes, sabiendo que la población de dicho establecimiento es de 200. Los datos obtenidos de *género*, *edad*, *rendimiento académico* (medido con un test estandarizado en: bajo, medio bajo, medio, medio alto y alto) y *horas de estudio por semana*, fueron los siguientes:

Sujeto	Género	Edad	Rendimiento académico	Hs. de estudio/semana
1	Femenino	16	Bajo	5
2	Masculino	16	Alto	10
3	Masculino	15	Medio	8
4	Femenino	17	Medio alto	9
5	Femenino	15	Medio	8
6	Femenino	15	Alto	10
7	Masculino	17	Medio alto	9
8	Masculino	14	Alto	12
9	Masculino	14	Medio bajo	7
10	Masculino	15	Medio bajo	7
11	Masculino	15	Medio	8
12	Femenino	13	Bajo	5
13	Masculino	15	Medio	8
14	Masculino	14	Medio bajo	7
15	Femenino	14	Medio alto	9
16	Femenino	16	Bajo	6
17	Masculino	15	Medio bajo	7
18	Masculino	14	Medio alto	9
19	Femenino	15	Alto	11
20	Masculino	17	Alto	11

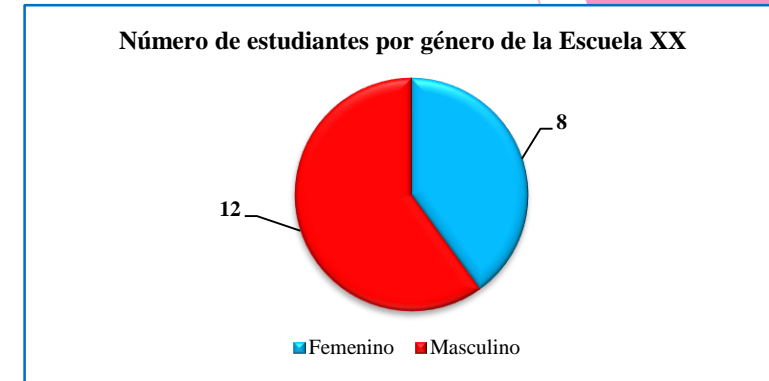


EJEMPLO (CONT.)

- **Género:** *Variable cualitativa – Nominal*

Estudiantes según género de la Escuela XX

Género	Nº de estudiantes	Proporción de estudiantes	% de estudiantes
Femenino	8	0,4	40
Masculino	12	0,6	60
	20	1	100



- **Edad:** *Variable cuantitativa continua – Métrica*

$$R = x_{\text{Max}} - x_{\text{Min}} = 17 - 13 = 4 \Rightarrow i = \frac{R}{\sqrt{n}} = \frac{4}{\sqrt{20}} = 0,894 \cong 1$$



Estudiantes según rango etario de la Escuela XX

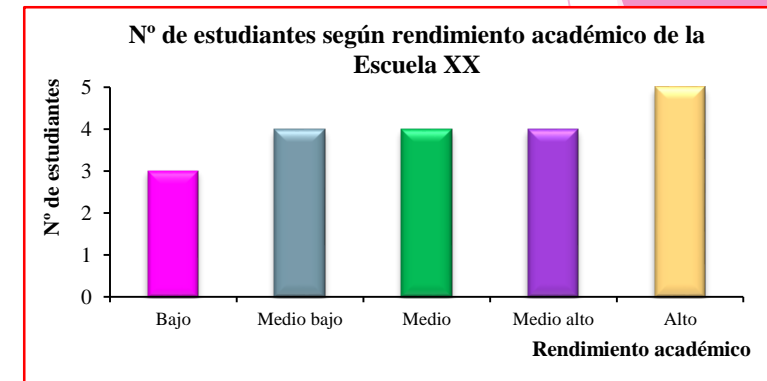
Rango etario	Pto medio del Rango etario	Nº de estudiantes	Proporción de estudiantes	% de estudiantes
13 - 14	13,5	1	0,05	5
14 - 15	14,5	5	0,25	25
15 - 16	15,5	8	0,40	40
16 - 17	16,5	3	0,15	15
17 - 18	17,5	3	0,15	15
		20	1	100



EJEMPLO (CONT.)

■ Rendimiento académico: *Variable cuantitativa continua – Ordinal* Estudiantes según rendimiento académico de la Escuela XX

Rendimiento académico	Nº de estudiantes	Proporción de estudiantes	% de estudiantes
Bajo	3	0,15	15
Medio bajo	4	0,20	20
Medio	4	0,20	20
Medio alto	4	0,20	20
Alto	5	0,25	25
	20	1	100



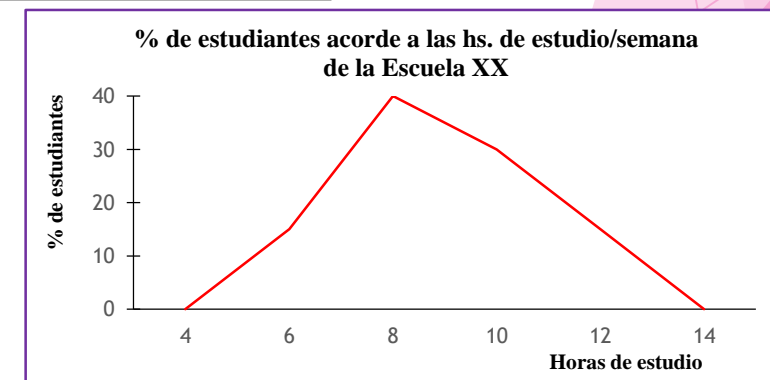
■ Horas de estudio: *Variable cuantitativa continua – Métrica*

$$R = X_{\text{Max}} - X_{\text{Min}} = 12 - 5 = 7 \Rightarrow i = \frac{R}{\sqrt{n}} = \frac{7}{\sqrt{20}} = 1,565 \cong 2$$



Estudiantes según hs. de estudio/semana de la Escuela XX

Hs. de estudio/semana	Punto medio	Nº de estudiantes	Proporción de estudiantes	% de estudiantes
5 - 7	6	3	0,15	15
7 - 9	8	8	0,40	40
9 - 11	10	6	0,30	30
11 - 13	12	3	0,15	15
		20	1	100



REFERENCIAS BIBLIOGRÁFICAS

- Bologna, E. (2011). *Estadística para Psicología y Educación*. Córdoba: Brujas.
- Gorgas García, J., Cardiel López, N. & Zamorano Calvo, J. (2009). *Estadística Básica para Estudiantes de Ciencias*. Madrid: Departamento de Astrofísica y Ciencias de la Atmósfera. Facultad de Ciencias Físicas. Universidad Complutense de Madrid.
- Penna, F.O., Esteva, G.C., Cobos, O.H. & Ulagnero, C.A. (2018). *Fórmulas y Tablas III (para cursos de Estadística básica)* (2ª ed.). San Luis: Nueva Editorial Universitaria.
- Sabulsky, J. (2000) *Investigación científica en salud-enfermedad* (3ª ed.) Córdoba: Ed. Kosmos.
- Triola M. (2018). *Estadística* (12ª ed.). México: Pearson Educación.