

LECTURA N°1

ESCALAS DE CLASIFICACIÓN

(Elaborado por Dra. Maritza Gil)

OBJETIVOS:

- Describir los métodos que pueden emplearse para el cómputo de los datos en una investigación.
- Identificar los tipos de variables
- Construir correctamente escalas de clasificación para ordenar series de datos.
- Determinar los límites verdaderos, amplitud, punto medio y frecuencia de cada clase de la escala.

Previo al análisis de los datos de una investigación, estos deben someterse a revisión rigurosa que permita corregir las inconsistencias ocurridas durante el proceso de recolección así como al ordenamiento necesario para realizar el análisis adecuado de los mismos y de este modo garantizar los resultados. A continuación se presentan los métodos utilizados para el conteo de los datos y los procedimientos utilizados para ordenarlos, antes de proceder a la presentación de los resultados. Se espera con ello, que el estudiante disponga de una herramienta de consulta rápida que le permita decidir sobre la forma más adecuada para manejar los datos de una investigación. Además de los planteamientos teóricos del proceso, se presentan ejemplos y ejercicios con los que pueda poner en práctica los conocimientos adquiridos.

CÓMPUTO DE LOS DATOS

Para analizar los datos recogidos en una investigación es necesario ordenarlos, pero previo a ello es conveniente hacer una revisión para detectar en primer lugar, si están completos los datos, si existieron problemas en la recolección y si es posible su corrección. Por ejemplo, en el caso de existir incongruencias entre la fecha de nacimiento y la edad de una persona o entre la edad y el peso, o que los datos estén incompletos, se debe realizar la corrección o decidir si los mismos se excluyen de los conteos, con lo cual se evitan errores en la presentación de los resultados.

Terminada la revisión se procede al cómputo de los datos para lo cual se pueden utilizar métodos manuales o mecánicos.

MÉTODOS MANUALES:

Método de las listas:

Según este procedimiento, las características estudiadas son asentadas en una lista que se denomina HOJA DE TRABAJO O SÁBANA DE DATOS. Un ejemplo se presenta a continuación:

EDAD (años)	SEXO	OCUPACION	EDO. CIVIL	CAUSA DE MUERTE
58	M	Agricultor	Casado	Infarto al miocardio
21	F	Oficios del hogar	Casada	Neumonía
13	F	Estudiante	Soltero	Dengue Hemorrágico
5	M	Sin ocupación	Soltero	Cardiopatía congénita
.
.
45	M	Mecánico	Soltero	Infarto al miocardio

Este método es útil cuando se trata de pocos datos y se realizan pocas combinaciones de variables.

Método de los palotes:

Partiendo de un listado como el presentado, o directamente del formulario de recolección de datos, se puede anotar en hoja aparte un palote cada vez que se encuentre una unidad que posea la(s) característica(s) a estudiar o las combinaciones de ella que se hagan.

Con este método se pueden cometer muchas equivocaciones si se trata de datos numerosos o cuando se combinan muchas características. El mayor problema lo representa la imposibilidad de verificar si un dato fue incluido o no, por lo cual ante cualquier duda o error, es necesario repetir todo el procedimiento.

Método de las tarjetas simples:

Consiste en preparar tarjetas de cartulina con los datos de cada unidad estudiada que incluya todas sus características. Permite trabajar con mayor número de datos que los procedimientos anteriores y realizar numerosas combinaciones de variables con pocas posibilidades de error y sobre todo con pocas dificultades para corregirlos. Por ejemplo, si con los datos anteriores se desea determinar el sexo de las personas, es fácil separar

las tarjetas por sexo y si además se quisiera separar por ocupación, se agrupa cada sexo por cada categoría.

EDAD: 58 años SEXO: M OCUPACION: Agricultor EDO. CIVIL: Casado CAUSA DE MUERTE: Infarto al Miocardio	EDAD: 45 años SEXO: M OCUPACION: Mecánico EDO. CIVIL: Soltero CAUSA DE MUERTE: Infarto al Miocardio	EDAD: 5 años SEXO: M OCUPACION: Sin ocupación EDO. CIVIL: Soltero CAUSA DE MUERTE: Cardiopatía congénita
--	---	--

Método de las tarjetas perforadas (perforaciones marginales):

Permite estudiar gran número de unidades y realizar todo tipo de combinaciones. Se prepara una tarjeta de cada unidad de estudio y en los márgenes de la misma se inscriben todas las características consideradas, sobre la cual se hace una perforación. Para señalar la característica que posee una unidad en particular, se rompe el borde de la tarjeta de modo que al pasar una punta sobre ese orificio caerán todas las unidades que la presenten.

Este método, al igual que el anterior, es poco utilizado por lo laborioso y por ventajas de los métodos computarizados.

Métodos automatizados:

Actualmente existen numerosos paquetes de computación (hojas de cálculo como Excel y paquetes estadísticos como EPI-INFO, SPSS, SAS, STATA, etc.) que permiten el procesamiento de los datos de una investigación. Con ellos se puede realizar cualquier combinación de variables (características) y obtener los resultados en muy poco tiempo cualquiera sea el número de unidades en estudio. Requiere la inclusión de los datos y la elaboración de una “base de datos” que debe ser revisada antes de proceder a los cómputos para corregir cualquier error cometido al incluirlos. Para su utilización se requiere preparación en el manejo del paquete y ofrecen la ventaja de permitir el análisis estadístico de los datos.

MÉTODOS DE COMPUTACION DE DATOS

1. Método de listas
2. Método de palotes
3. Método de tarjetas simples
4. Método de tarjetas perforadas
5. Métodos automatizados

TIPOS DE VARIABLES

Es importante saber qué tipo de variables estamos estudiando pues de ello depende el tipo de gráficos que se podrá emplear para representar los datos y las medidas que podremos calcular para resumirlos.

- a) **Variables cualitativas:** son aquellas que expresan cualidades o características de la unidad en estudio, por ejemplo sexo, estado civil, ocupación, etc.
- b) **Variables cuantitativas:** son aquellas que, además de evidenciar las características, permiten su medición (cuantificación) por ejemplo peso, edad, cifras de hemoglobina, etc.

Las variables cuantitativas pueden ser continuas y discontinuas. Se habla de variable **cuantitativa continua** cuando el valor numérico admite fracciones, por ejemplo un individuo puede pesar 85 kilos con 600 gramos, mientras que la variable **cuantitativa discreta** (o discontinua) no permite fracciones, admitiendo solamente números enteros, por ejemplo el número de caries de un sujeto.

ESCALAS DE CLASIFICACIÓN

Como se ha mencionado, luego de revisar los datos de la investigación, es necesario ordenarlos para que puedan ser analizados. Para ello deben elaborarse las escalas de clasificación que servirán para ubicar correctamente cada unidad de investigación. Si por ejemplo, tenemos las edades de un grupo de 35 personas:

5	19	21	18	22	27	12
16	20	18	13	25	14	17
9	20	14	5	17	12	19
11	10	23	17	21	26	10
19	28	16	21	25	11	19

es difícil interpretar la información a menos que elaboremos una escala que facilite su análisis:

Edad (años)	No. de personas
5 - 9	3
10-14	9
15-19	11
20-24	7
25-29	5
TOTAL	35

Al construir una escala de clasificación debemos tomar en consideración que cumpla con dos condiciones:

- Debe ser **exhaustiva**, es decir, que permita la clasificación de todos los individuos incluidos en el estudio.
- Que las clases sean **mutuamente excluyentes** lo cual no deja lugar a dudas sobre la clase en la cual se debe incluir cada unidad estudiada.

A continuación se presentan otros ejemplos con los datos anteriores:

INCORRECTA	CORRECTA	INCORRECTA	CORRECTA
5 - 8	5 - 8	0 - 5	0 - 5
8 - 12	9 - 12	5 -11	6 -11
12 - 16	13 - 16	12 -17	12 -17
16 - 20	17 - 20	17 -23	18 -23
20 - 24	21 - 24	23 -29	24 -29
24 - 28	25 - 28		

CARACTERISTICAS DE UNA BUENA ESCALA

1. Exhaustiva
2. Mutuamente excluyente

Por otra parte, al momento de construir una escala, es necesario establecer la amplitud de las clases a estudiar puesto que no es conveniente elaborarlas con muy pocas clases pues algunas diferencias importantes podrían pasar inadvertidas o con tantas clases que diluyan las características del grupo. Dependiendo del estudio, se recomienda elaborar escalas entre 5 y 20 clases.

Tomando los datos del ejemplo anterior, se presentan a continuación las diferencias al construir escalas con muy pocas o demasiadas clases.

ESCALA CON MUY POCAS CLASES

<u>EDAD (años)</u>	<u>No.</u>
--------------------	------------

Al utilizar una escala inadecuada, la mayoría de los sujetos quedan ubicados en una sola clase

5 - 14	12	
15 - 24	18	←
25 - 35	5	
<hr/>		
	n = 35	

ESCALA CON MUCHAS CLASES

EDAD (años)	No.
5 - 6	2
7 - 8	0
9 - 10	3
11 - 12	4
13 - 14	3
15 - 16	2
17 - 18	5
19 - 20	6
21 - 22	4
23 - 24	1
25 - 26	3
27 - 28	2
<hr/>	
	n = 35

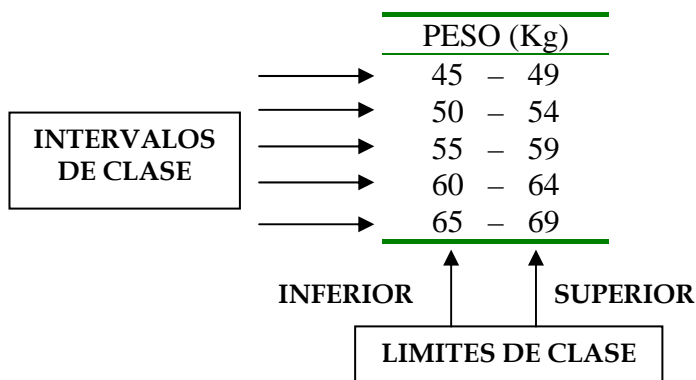
Nótese como la información se diluye y las diferencias entre las clases son muy poco evidentes.

PROCEDIMIENTO PARA LA CONSTRUCCION DE UNA ESCALA

Al construir una escala es necesario considerar varios aspectos a fin de no cometer errores que puedan afectar la interpretación de los resultados.

Intervalos y límites de clase:

En primer lugar se debe saber que el intervalo de una clase es el símbolo que la define, 45-49 y 50-54, por ejemplo, son intervalos de clase mientras que los números de los extremos son los límites de la clase.



Cada intervalo tiene un límite inferior (LI) y un límite superior (LS). De forma general puede presentarse de la siguiente manera:

LI	LS	frecuencia
----	----	------------

L_1	L_2	f_1
L_2	L_3	f_2
L_3	L_4	f_3
L_n	L_{n+1}	f_n

Sin embargo, esos son límites aparentes pues a la hora de incluir las unidades en estudio, se deben considerar los límites verdaderos. Si tenemos un individuo que pesa 49,5 Kg, podríamos incluirlo en la clase de 45-49 si deseamos la fracción, o en la clase de 50-54 si redondeamos la cifra.

Los límites verdaderos se obtienen restando al límite aparente inferior la mitad de la unidad de medida. En nuestro ejemplo, como la unidad de medida es 1 Kg, la mitad es 0,5. Así, 45 menos 0,5 = 44,5. La escala queda de la siguiente forma:

<u>LIMITES APARENTES</u>	<u>LIMITES VERDADEROS</u>
45 – 49	44,5 – 49,5
50 – 54	49,5 – 54,5
55 – 59	54,5 – 59,5
60 – 64	59,5 – 64,5
65 – 69	64,5 – 69,5

A pesar del procedimiento anterior, aún podrían presentarse dudas para incluir al individuo que pesó 49,5 Kg. Es por ello que al límite verdadero superior se le resta una fracción decimal equivalente al número de dígitos considerado, en el ejemplo la fracción es igual a 0,01, por lo cual los límites verdaderos son los siguientes:

<u>LIMITES APARENTES</u>	<u>LIMITES VERDADEROS</u>
45 – 49	44,5 – 49,49
50 – 54	49,5 – 54,49
55 – 59	54,5 – 59,49
60 – 64	59,5 – 64,49
65 – 69	64,5 – 69,49

Lo anteriormente expuesto se cumple en todos los casos excepto con la edad pues en ese caso hay que considerar que la edad no se aproxima al cumpleaños siguiente si no que se cuentan los años cumplidos, de este modo una persona que tiene 25 años tendrá esa edad hasta el momento de cumplir 26 años, aún cuando para ello le falte un día. La misma consideración se debe tener en cuenta respecto al límite inferior pues una vez que se alcanza una edad, no se tiene menos de esa edad.

LIMITES APARENTES

5 - 9
 10 - 14
 15 - 19
 20 - 24
 25 - 29

LIMITES VERDADEROS

5 - 9,99
 10 - 14,99
 15 - 19,99
 20 - 24,99
 25 - 29,99

Amplitud de la clase:

Es la diferencia entre los valores máximo y mínimo de una clase (entre los límites verdaderos inferior y superior). La amplitud de las clases puede ser la misma para todas las clases o pueden construirse clases con amplitudes diferentes, lo último es frecuente en los estudios epidemiológicos donde suelen construirse clases de edad con amplitud diferente tomando en cuenta los grupos de riesgo, por ejemplo:

EDAD (años)
Menores de 1
1 - 4
5 - 9
10 - 14
15 - 24
25 - 44
45 - 54
55 - 64
65 y más

A este ejemplo la amplitud de las clases sería:

EDAD	LIMITES VERDADEROS	AMPLITUD (LVS - LVI)
Menores de 1	0 - 0,99	0,99 \cong 1
1 - 4	1 - 4,99	3,99 \cong 4
5 - 9	5 - 9,99	4,99 \cong 5
10 - 14	10 - 14,99	4,99 \cong 5
15 - 24	15 - 24,99	9,99 \cong 10
25 - 44	25 - 44,99	19,99 \cong 20
45 - 54	45 - 54,99	9,99 \cong 10
55 - 64	55 - 64,99	9,99 \cong 10
65 y más	65 -	No se puede calcular

Punto medio (x_i):

Resulta de la semisuma de los límites verdaderos de una clase:

$$PUNTO\ MEDIO = \frac{LVS + LVI}{2}$$

Así, para la clase 50 – 54 kilogramos:

$$x_i = \frac{49,5 + 54,49}{2} ; x_i = 51,995$$

Frecuencia de clase (f_i) :

Viene dada por la cantidad de individuos estudiados que deben ser incluidos en cada clase. Ejemplo:

Para los datos sobre peso que se presentan a continuación se calcularon 5 clases, los límites verdaderos, el punto medio, la frecuencia y la amplitud de las mismas.

PESO EN KILOGRAMOS DE 25 PERSONAS

55	63	72	60	55
57	71	48	48	56
49	54	47	65	60
50	46	67	61	63
68	63	59	60	68

PESO (Kg)	LVI - LVS	FRECUENCIA	PUNTO MEDIO
46 - 50	45,50 - 50,49	6	47,995
51 - 55	50,50 - 55,49	3	52,995
56 - 60	55,50 - 60,49	6	57,995
61 - 65	60,50 - 65,49	5	62,995
66 - 70	65,50 - 70,49	3	67,995
71 - 75	70,50 - 75,49	2	72,995

amplitud de las clases = 5

tipo de variable = cuantitativa continua

PASOS PARA LA CONSTRUCCION DE ESCALAS DE CLASIFICACION

1. Establecer el número deseado de clases: se recomienda no construir escalas con menos de 5 clases ni mayores de 20 clases
2. Identificar el menor y mayor valor entre los datos recogidos y la diferencia entre ellos; ésta será la amplitud total de la serie de datos.
3. Dividir la amplitud de la serie de datos entre el número deseado de clases, de esta forma obtenemos la amplitud que debe tener cada una de las clases de la escala.

4. Determinar el valor por el cual se comenzará la construcción (menor número de la serie de datos o un número cercano). Ese valor representa el límite inferior de la primera clase. Luego se suma a ese valor la amplitud de la clase; así obtenemos el límite inferior de la segunda clase. Continuamos el mismo procedimiento hasta obtener los límites inferiores de todas las clases. Posteriormente restamos una unidad a estos límites para obtener los límites superiores de cada clase (por ejemplo, si el límite inferior de la segunda clase es 60, entonces el límite superior de la primera clase será 59). **EJEMPLO:**

PESO (Kg) DE 20 INDIVIDUOS			
50	59,6	63	75
50,4	60	64	80
50,6	61	65	82
50,6	62	68	90
50,6	62	70	97

1. Número deseado de clases = 5
2. Diferencia entre el mayor y menor valor: $97 - 50 = 47$
3. Amplitud entre número de clases: $47/5 = 9,4 \cong 10$
4. Valor por el cual se comenzará la construcción (menor número de la serie) = 50

PESO (Kg)	LVI - LVS	FRECUENCIA
50 - 59	49,50 - 59,49	5
60 - 69	59,50 - 69,49	9
70 - 79	69,50 - 79,49	2
80 - 89	79,50 - 89,49	2
90 - 99	89,50 - 99,49	2

En otro orden de ideas, las variables pueden estudiarse de varias formas:

1. **Distribución de frecuencias:** cuando se clasifica los datos según una única escala. Ejemplo, la distribución de los sujetos estudiados según su peso, su estatura, o su sexo.
2. **Datos de asociación:** cuando se combinan 2 o más variables para contrastar las diferencias existentes entre grupos. Ejemplo, la distribución de los sujetos del estudio, según su edad y sexo.
3. **Series cronológicas:** cuando se presenta la evolución de un fenómeno en el tiempo.

EJEMPLO DE DISTRIBUCIÓN DE FRECUENCIAS

PESO (Kg)	FRECUENCIA
46 - 50	6
51 - 55	3
56 - 60	6
61 - 65	5
66 - 70	3
71 - 75	2

EJEMPLO DE DATOS DE ASOCIACIÓN

PESO (Kg)	SEXO	
	MASCULINO	FEMENINO
46 - 50	5	9
51 - 55	7	12
56 - 60	6	11
61 - 65	9	8
66 - 70	11	5
71 - 75	10	2

EJEMPLO DE SERIES CRONOLÓGICAS

AÑO	FRECUENCIA
1993	63
1994	76
1995	55
1996	73

EJERCICIOS

1. Con los datos que se presentan a continuación que se refieren a la edad en años de una muestra de personas, construya una escala de clasificación de amplitud 5 en cada clase; calcule los límites verdaderos, frecuencia y punto medio. Diga qué tipo de variable es la estudiada en ese ejemplo.

15	26	21	23	10
12	32	27	12	21
23	23	20	23	23
21	33	19	11	19
18	35	17	21	18
11	21	35	33	17

2. Con los datos del ejercicio anterior, construya una escala de clasificación de 5 clases, calcule los límites verdaderos, frecuencia y punto medio

3. Para los enunciados que se presentan a continuación, mencione el tipo de variable a que se refiere cada uno de ellos:
 - a) Clasificación de los individuos según niveles de colesterol
 - b) Tipo de tratamiento recibido por los pacientes que acuden a un hospital
 - c) Mujeres según número de hijos
 - d) Niños según número de caries
 - e) Tamaño del grupo familiar

4. Se presenta a continuación el promedio de calificaciones de un grupo de estudiantes. Con los datos, construya una escala de clasificación de cinco clases (comenzar en 10). Calcule los límites verdaderos, frecuencia y punto medio. Indique el tipo de variable.

11,51	16,82	14,28	15,27	10,93
12,36	13,49	11,75	10,49	12,16
12,34	11,35	12,09	12,39	16,73
10,61	13,23	19,13	11,18	16,52
18,12	13,54	17,21	12,49	15,68
11,29	12,15	13,55	13,39	17,25

BIBLIOGRAFIA CONSULTADA

1. Ary D, Jacobs A, Razavieh A. Introducción a la Investigación Pedagógica. México: Interamericana; 1986

2. Camel F. Estadística Médica y de Salud Pública. Tomo I. Mérida, Venezuela.: Ediciones Universidad de los Andes; 1988

3. Chourio J. Estadística I. Caracas: Editorial Biosfera; 1987

4. Haber A, Runyon R. Estadística General. USA: Fondo Educativo Interamericano; 1973

5. Levin J. Estadística en la Investigación Social. 2ª. Edición. México: Harper & Row Latinoamericana. 1979
6. Polit D, Hundler B. Investigación Científica en Ciencias de la Salud. 4ta. Edición. México: McGraw Hill. 1994.