

SESIÓN 2

ESTADÍSTICA DESCRIPTIVA

I. CONTENIDOS:

1. Tablas estadísticas.
2. Distribución de frecuencias para datos no agrupados.
3. Distribución de frecuencias para datos agrupados.
4. Terminología relativa a datos agrupados.

II. OBJETIVOS:

Al término de la Sesión, el alumno:

- Realizará tablas con todos los tipos de variables, con datos agrupados y datos no agrupados.
- Creará tablas con datos agrupados en intervalos de clases para variables cardinales.

III. PROBLEMATIZACIÓN:

Comenta las preguntas con tu Asesor y selecciona las ideas más significativas.

- ¿Cómo representarías los datos obtenidos de una encuesta para su estudio y mejor comprensión?
- ¿Cuales son las ventajas de ordenar un conjunto de datos para su estudio?
- ¿Conoces algún caso en tu comunidad que quisieras investigar mediante datos estadísticos?

IV. TEXTO INFORMATIVO-FORMATIVO:

1.1. Tablas estadísticas

La presentación de datos que han sido recolectados, analizados y sintetizados es uno de los más importantes productos que ofrece la estadística. Los resultados del trabajo estadístico se muestran en tablas o elementos gráficos.

Una tabla estadística es una descripción de las frecuencias de las categorías de una variable, a partir de ella se elaboran los gráficos. Los elementos que la constituyen son:

1. *El encabezamiento: es la parte superior, en ella aparece una descripción del contenido de la tabla mediante un título alusivo, debajo se encuentra el periodo que es el tiempo en que se realizó el estudio, además las unidades de medición si hace falta.*
2. *El cuerpo: es la parte más importante, en ella se enumeran las categorías de la variable y sus respectivas frecuencias. Se disponen a manera de columnas de una manera lógica para ser leídas de izquierda a derecha.*
3. *El pie: en esta parte se agregan notas aclaratorias y la fuente de donde se extrajo la información.*

A continuación se muestra un ejemplo de una tabla estadística. En ella se puede apreciar:

1. El encabezamiento: "Años de servicio de los docentes (n=89)"
2. El cuerpo: con las categorías "PROGRAMA" y las frecuencias para diferentes rangos de "AÑOS DE SERVICIO"
3. El pie: "Cuestionarios aplicados a docentes"

TABLA 7. Años de servicio de los docentes (n = 89)

PROGRAMA	AÑOS DE SERVICIO				
	0 a 4	5 a 9	10 a 14	15 a 19	20 o más
MDE	3	10	19	17	6
MPE	0	3	5	4	1
MP	1	0	16	3	1
TOTAL	4	13	40	24	10

Fuente: Cuestionarios aplicados a docentes.

Consultado el 30 de Marzo de 2011 de http://www.scielo.unal.edu.co/scielo.php?pid=S0121-50512009000400012&script=sci_arttext

2.1. Distribución de frecuencias de datos no agrupados

Toda variable que ha sido estudiada, se describe en sus diferentes categorías de una manera que facilite el estudio de la manera como se manifiesta. Las categorías de toda variable deben ser:

- Mutualmente excluyentes: un dato observado no puede asociarse más que a una de las categorías.*
- Exhaustivas: cada dato observado debe asociarse con alguna categoría.*

La descripción de las categorías de una variable, de manera ordenada y precisa, se conoce como **distribución de frecuencias**. Donde la **frecuencia** es el número de repeticiones observadas de cada categoría. Para las variables nominales sólo es posible describir la frecuencia de cada categoría y su correspondiente frecuencia relativa o porcentaje. En cambio, en las variables ordinales, además de esto, se puede conocer la categoría que queda en el centro de la distribución y también la llamada frecuencia acumulada. Para construir la distribución de frecuencias acumuladas, se procede de la siguiente manera:

- Se colocan en una columna las categorías de la variable de acuerdo a su jerarquía.
- A la derecha de agrega otra columna con las frecuencias asociadas a las categorías.
- Se suman las frecuencias y el resultado debe coincidir con el número de datos.
- Se divide cada frecuencia entre el número de datos y el resultado se multiplica por cien, estos valores forman una nueva columna llamada de frecuencias relativas.
- Se suman progresivamente las últimas dos columnas, una a la vez, y se obtienen otras dos columnas llamadas de frecuencias acumuladas ascendentes.
- Se restan desde el número de datos, de manera progresiva, las columnas de las frecuencias y frecuencias relativas y se obtienen otras dos columnas llamadas de frecuencias acumuladas descendentes.
- Se agrega título a la tabla, periodo y fuente.

Los siguientes ejemplos pueden clarificar el proceso:

Ejemplo 1. Se entrevistaron a 30 personas en una concurrida tienda departamental de Puerto Vallarta, julio de 2004, a fin de conocer si eran turistas y en dado caso registrar el Estado de la república de donde procedían. Se obtuvieron los siguientes datos:

Jalisco, Jalisco, Colima, Michoacán, Jalisco, Nayarit, Nayarit, Jalisco, Colima, Michoacán, Guanajuato, Jalisco, Zacatecas, Michoacán, Aguascalientes, Estado de México, San Luis Potosí, Jalisco, Colima, Michoacán, Guanajuato, Jalisco, Guanajuato, Zacatecas, Aguascalientes, Nayarit, Nayarit, Colima, Jalisco, Nayarit. Elabora una distribución de frecuencias con la información.

Para elaborar la distribución de frecuencias se disponen en una columna las categorías y a la derecha el número de veces que se repiten, luego se suman para comparar con el número de datos:

Estado	Frecuencia
Jalisco	8
Colima	4
Michoacán	4
Nayarit	5
Guanajuato	3
Zacatecas	2
Aguascalientes	2
Estado de México	1
San Luis Potosí	1
Suma	30

Enseguida, se divide cada frecuencia entre el número de datos y al cociente se le multiplica por cien para obtener un porcentaje:

$8 \div 30 \times 100 =$	26.66 %
$4 \div 30 \times 100 =$	13.33 %
$4 \div 30 \times 100 =$	13.33 %
$5 \div 30 \times 100 =$	16.66 %
$3 \div 30 \times 100 =$	10 %
$2 \div 30 \times 100 =$	6.66 %
$2 \div 30 \times 100 =$	6.66 %
$1 \div 30 \times 100 =$	3.33 %
$1 \div 30 \times 100 =$	3.33 %

Finalmente se dispone la información en una tabla de distribución de frecuencias:

Turismo nacional en una tienda departamental en Puerto Vallarta

Julio de 2004

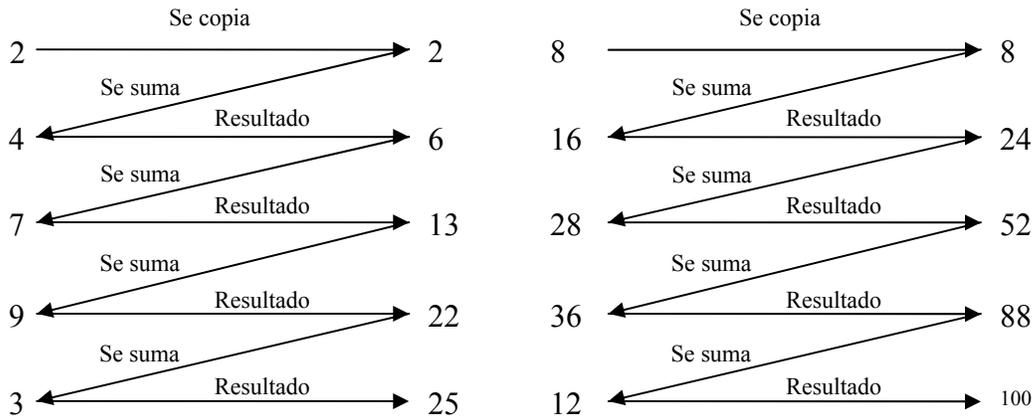
Estado de la República	Frecuencia 30	Frecuencia relativa %
Jalisco	8	26.66
Colima	4	13.33
Michoacán	4	13.33
Nayarit	5	16.66
Guanajuato	3	10
Zacatecas	2	6.66
Aguascalientes	2	6.66
Estado de México	1	3.33
San Luis Potosí	1	3.33

Ejemplo 2 Se registraron las evaluaciones finales de un grupo de 25 trabajadores, después de un curso de capacitación, en la empresa “Hilos de México S.A.”, febrero de 2001, obteniendo los siguientes resultados: Excelente, bien, muy bien, regular, regular, bien, bien, regular, no aprobado, no aprobado, regular, bien, excelente, regular, muy bien, regular, bien, bien, regular, bien, muy bien, no aprobado, muy bien, regular, regular. Elabora una distribución de frecuencias acumuladas.

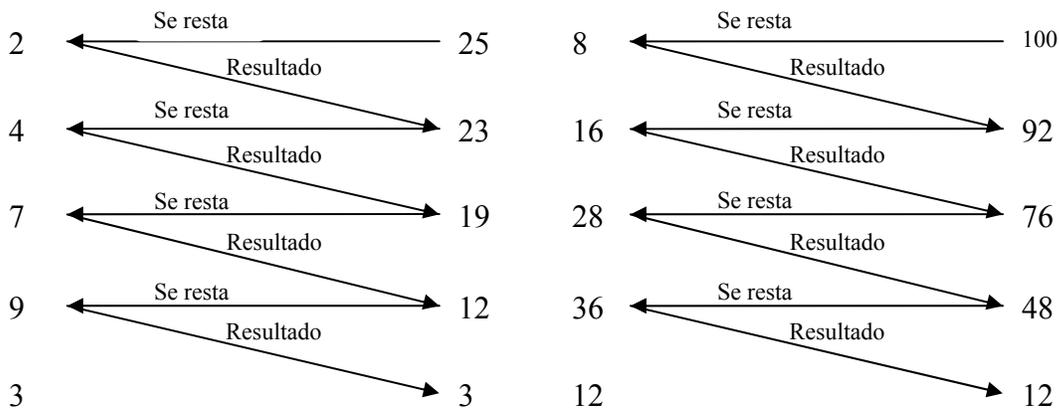
Se organizan las categorías de manera jerárquica y se procede como en el ejemplo anterior hasta tener una tabla como la siguiente:

Evaluación	Frecuencia 25	Frecuencia relativa %
Excelente	2	8
Muy bien	4	16
Bien	7	28
Regular	9	36
No aprobado	3	12

Después se suman de manera progresiva las frecuencias, para encontrar las frecuencias acumuladas ascendentes f. a. (+) y f. a. (+) %



Luego, se copia el total de datos y se resta progresivamente de las frecuencias, para encontrar las f. a. (-) y f. a. (-) %



Ahora se dispone la distribución de frecuencias

Evaluación	f	f. r %	f. a. (+)	f. a. (+) %	f. a. (-)	f. a. (-) %
Excelente	2	8	2	8	25	100
Muy bien	4	16	6	24	23	92
Bien	7	28	13	52	19	76
Regular	9	36	22	88	12	48
No aprobado	3	12	25	100	3	12

Vamos a leer un poco de esta tabla:

- El número 13 de la cuarta columna, significa que 13 trabajadores obtuvieron una evaluación al menos de “Bien”, representando el 52% del total.
- El número 76 de la séptima columna significa que 19 trabajadores, que representan el 76%, obtuvieron una evaluación de “Bien” o menos.
- El número 36 de la tercera columna significa que nueve trabajadores, que representan el 36%, obtuvieron una evaluación de “Regular”

En las variables cardinales discretas se puede desarrollar una distribución de frecuencias semejante, aunque como las categorías son números se integran formando otras columnas como veremos mas adelante.

3.1. Distribución de frecuencias para datos agrupados

Si se trata de una variable cardinal continua, y el número de observaciones es grande, se recomienda construir una distribución de datos agrupados. El procedimiento más común es:

1. Se ordenan de menor a mayor las todas las observaciones.
2. Se resta del dato mayor el dato menor, el resultado es el rango.
3. Para que no coincidan los datos con los límites de los intervalos se suma al rango 1 si los datos son enteros, 0.1 si los datos tienen una cifra decimal, 0.01 si dos, etc.
4. Se determina el número de intervalos, se recomienda que sea entre 5 y 20.
5. Se calcula la amplitud de clase dividiendo el resultado del paso 3 entre el número de intervalos.
6. Determinar el Límite Real Inferior (LRI) del primer intervalo, restando al dato menor la mitad de lo que se sumó al rango en el paso 3.
7. Desarrollar los LRI y LRS de cada intervalo sumando progresivamente la amplitud.
8. Contar los datos que se encuentran entre los LRI y LRS de cada intervalo y registrarlos como frecuencia.
9. Proceder como en el ejemplo 2 para integrar las frecuencias acumuladas.

Ejemplo 3 Los siguientes valores representan los tiempos, en segundos, que tardaron 54 operarios de una ensambladora de computadoras, en montar la tarjeta madre en una computadora de escritorio. Construye una distribución de frecuencias, tanto relativas como acumuladas, de 9 intervalos de amplitud constante.

345	626	688	506	700	240	363	670	396
474	448	670	1023	335	540	853	495	666
483	443	325	514	404	820	915	778	627
552	991	615	609	723	794	447	396	235
748	756	587	929	871	378	503	1128	693
862	861	337	292	1070	625	457	494	420

Ordenando los datos de menor a mayor:

235	240	292	325	335	337	345	363	378
396	396	404	420	443	447	448	457	474
483	494	495	503	506	514	540	552	587
609	615	625	626	627	666	670	670	688
693	700	723	748	756	778	794	820	853
861	862	871	915	929	991	1023	1070	1128

Se calcula el rango: $1128 - 235 = 893 + 1 = 894$

Amplitud de clase = $894 \div 9 = 99.33$

LRI del primer intervalo = $235 - 0.5 = 234.5$

LRI	LRS	f	f. r. %	f. a. (+)	f. a. (+) %
234.5	333.83	4	7.40	4	7.40
333.83	433.16	9	16.66	13	24.06
433.16	532.49	11	20.37	24	44.43
532.49	631.82	8	14.81	32	59.24
631.82	731.15	7	12.96	39	72.2
731.15	830.48	5	9.25	44	81.45
830.48	929.81	6	11.11	50	92.56
929.81	1029.14	2	3.70	52	96.26
1029.14	1128.47	2	3.70	54	100

V. ESTRATEGIAS CENTRADAS EN EL APRENDIZAJE

Construye una distribución de frecuencias con 12 intervalos de los siguientes datos que representan el espesor de latas de refresco en pulgadas.

.1009 .0097 .0098 .0113 .0110 .0103 .0095 .0104 .0101
 .0113 .0096 .0113 .0112 .0095 .0104 .0104 .0102 .0103
 .0101 .0098 .0100 .0099 .0107 .009 .0113 .0093 .0112
 .0099 .0100 .0101 .0098 .0106 .0097 .0111 .0099 .0115
 .0107 .0103 .0103 .0111 .0105 .0096 .0110 .0097 .0114
 .0105 .0104 .0104 .0103 .0103 .0015 .0105 .0107 .0115
 .0111 .0113 .0113 .0108 .0094 .0100 .0113 .0104 .0110
 .0110 .0112 .0113 .0107 .0098 .0114 .0112 .0096 .0106
 .0102 .0111 .0111 .0106 .0103 .0105 .0100 .0096 .0105
 .0101 .0106 .0109 .0097 .0113 .0107 .0093 .0113 .0107