



UNIVERSIDAD AUTÓNOMA DE CENTRO AMÉRICA
SEDE REGIONAL PACÍFICO NORTE
NICOYA - GUANACASTE

DISTRIBUCIONES DE FRECUENCIAS
DE VARIABLES DISCRETAS

No HERMANOS	FREC. ABS. (f_j)	FREC. REL. (f_j/n)	FRECUENCIA ACUMULADA "MENOS DE" (F_j)		FRECUENCIA ACUMULADA "MÁS DE" (F_j)	
			ABSOLUTA	RELATIVA	ABSOLUTA	RELATIVA
1	1	0,08	1	0,08	25	1,00
2	2	0,12	3	0,20	23	0,92
3	3	0,16	6	0,36	20	0,80
4	4	0,32	10	0,68	16	0,64
5	5	0,12	15	0,80	11	0,32
6	6	0,04	19	0,84	5	0,20
7	7	0,08	23	0,92	4	0,16
8	8	0,08	25	1,00	3	0,08
9	9	1,00	25	1,00	0	0,00

CURSO:
PROBABILIDAD Y ESTADÍSTICA

PREPARADO POR:
ING. ALLAN VILLEGAS ALEMÁN

LA NECESIDAD DE RESUMIR NUESTRA INFORMACIÓN

Los datos estadísticos provienen de la operación de contar o de la de medir; y pueden haberse obtenido anotando el número de elementos que corresponden a cada una de las categorías definidas o también ser el fruto de una operación más compleja, como la medición de la intensidad o magnitud de una característica.

Esto da origen, como se expuso anteriormente, a dos tipos de variables: cualitativas (o *atributos*) y cuantitativas (*distinguiéndose dentro de estas últimas las continuas y las discretas*).

En las exposiciones anteriores se vio la clasificación de datos, así como los requisitos mínimos y las reglas a cumplir para hacerlo bien. Se explicó cómo construir y representar las series cualitativas (*mediante cuadros y gráficos*). Ahora nos enfocaremos en la construcción de distribuciones de frecuencias.

Las distribuciones de frecuencias son clasificaciones que se refieren a variables cuantitativas, continuas o discretas y que constituyen un instrumento muy útil en el trabajo estadístico.

En el análisis e interpretación de los datos estadísticos correspondientes a variables continuas y discretas, resulta muy valioso disponer de elementos descriptivos que den información acerca de tres aspectos:

- a) La **forma** o patrón de la distribución de los datos.
- b) La **posición** de la distribución, o sea, alrededor de qué valor se tienden a concentrar los datos (*valores centrales*).
- c) La **dispersión** de los datos alrededor de los valores centrales o promedios (*variabilidad*).

Esta información puede lograrse fácilmente cuando el conjunto de interés lo forman pocos datos; basta ordenarlos de acuerdo a su magnitud y con un simple examen se podría conocer cuál es la forma de su distribución (a); el valor mayor y el menor, y la amplitud general (o sea, la diferencia entre esos dos valores extremos) que nos da una idea de su variabilidad (c); la existencia de concentraciones alrededor de algunos valores (b); además de otros detalles.

Cuando los datos son relativamente numerosos, sin embargo, lo anterior resulta insuficiente y se hace difícil apreciar las regularidades que existen en el conjunto, realizar análisis o sacar conclusiones; por ello debe recurrirse a agrupar los datos en una distribución de frecuencias, que puede definirse como *una ordenación o arreglo de datos en clases o categorías que muestran, para cada una de ellas, el número de elementos que contiene o frecuencia*.

Como hemos visto hasta el momento, existen dos tipos de variables: *continuas* y *discretas*. Antes bien, primero definiremos algunos conceptos necesarios para entender distribuciones de frecuencias de variable discreta.

Definiciones

Distribución (o Tabla) de Frecuencias: Es la representación conjunta de los datos en forma de tabla, correspondientes a un fenómeno en estudio y su ordenamiento es, con base al número de observaciones que corresponden a cada dato.

Se elabora colocando en la primera columna los datos diferentes y en la columna siguiente el número de observaciones que corresponden a cada dato llamada frecuencia. Una tabla de este tipo dará, en forma abreviada, una información completa acerca de la distribución de los valores observados. Estas tablas facilitan el uso de los métodos gráficos y aritméticos.

La presentación de los datos en forma ordenada, por medio de una tabla, dependerá de los datos de que se trate, y si estos son cualitativos o cuantitativos como se muestra a continuación:

Datos	Ordenamiento
Cualitativos	Alfabético A – Z
	Alfabético Z – A
	Del más al menos repetido
	Del menos al más repetido
Cuantitativos	Creciente (del menor al mayor)
	Decreciente (del mayor al menor)

EJEMPLO 1: Cierta universidad realizó un experimento sobre el coeficiente intelectual (C.I.) de sus alumnos, para lo cual aplicó un examen de C.I. a un grupo de 20 alumnos escogidos al azar, obteniendo los siguientes resultados:

119, 109, 124, 119, 106, 112, 112, 112, 112, 109, 112, 124, 109, 109, 109, 106, 124, 112, 112, 106.

Toda vez que se tienen los datos, se ordenan de menor a mayor o viceversa.

106, 106, 106, 109, 109, 109, 109, 109, 112, 112, 112, 112, 112, 112, 112, 112, 119, 119, 124, 124, 124.

Tabla 1: Distribución de frecuencias para el ejemplo1.

DATO	REPETICIONES (Frecuencia con que se repite el dato)
106	3
109	5
112	7
119	2
124	3

Frecuencia Absoluta (de un dato): Es el número de veces que se repite ése dato. La denotaremos por f_i .

Frecuencia Absoluta Acumulada (de un dato): Hasta un dato específico, es la suma de las frecuencias absolutas de todos los datos anteriores, incluyendo también la del dato mismo del cual se desea su frecuencia acumulada. La última frecuencia absoluta acumulada deberá ser igual al número total de datos. La denotaremos por f_a .

Frecuencia Relativa (de un dato): Se obtiene al dividir la frecuencia absoluta de cada dato entre el número total de datos. La denotamos por f_r .

Frecuencia Relativa Acumulada (de un dato): Hasta un dato específico de la observación, es la suma de las frecuencias relativas de todos los datos anteriores, incluyendo también la del dato mismo del cual se desea su frecuencia relativa acumulada. La última frecuencia relativa acumulada deberá ser igual a la unidad. La denotaremos por f_{ra} .

DISTRIBUCIONES DE FRECUENCIAS DE VARIABLES DISCRETAS

Para ilustrar la construcción de distribuciones de frecuencias para variables discretas, considérese el siguiente ejemplo.

EJEMPLO 2: A usted le contratan en el departamento de “Becas y Estímulos” de un centro de estudios superiores. A usted como profesional en psicología se le encomienda trabajar en conjunto con un profesional en trabajo social. A ambos se les asignó un grupo de 25 alumnos de primer ingreso a la universidad. La idea es que evalúen la posibilidad de otorgarles una ayuda socioeconómica a los estudiantes que así más lo requieran.

El primer paso que ustedes dieron fue, anotar cuántos hermanos tiene cada uno de los 25 estudiantes. Lo que dio como resultado la siguiente lista:

ESTUDIANTE	Nº HERMANOS	ESTUDIANTE	Nº HERMANOS	ESTUDIANTE	Nº HERMANOS
AA	3	JB	2	MA	3
RA	5	CA	4	RC	4
AD	5	JE	7	JE	6
MM	8	LM	4	LP	2
RQ	7	PR	4	CV	8
RG	3	GG	1	AG	3
FG	2	HH	4	VH	4
RJ	4	MJ	1	MV	5
JZ	4				

Un análisis más cómodo y eficiente puede lograrse agrupando las respuestas en una distribución, tal como se indica abajo.

Tabla 2: Distribuciones de frecuencias para el ejemplo 2.

Nº HERMANOS (1)	RECUENTO (2)	FRECUENCIA ABSOLUTA (f_i) (3)	FRECUENCIA RELATIVA (f_r) (4)
1	//	2	0,08
2	///	3	0,12
3	////	4	0,16
4	#### ///	8	0,32
5	///	3	0,12
6	/	1	0,04
7	//	2	0,08
8	//	2	0,08
		25	1,00

Note que la frecuencia para cada número de hermanos se logra marcando una línea diagonal (/) cada vez que aparece una observación que corresponde a ése número (*columna 2*). Nótese también que para facilitar la totalización posterior, cada quinta observación se marca con una línea en medio (*como una tacha*) que cruza las anteriores.

La distribución muestra (*columna 3*) que el número más frecuente de hermanos es 4 y que la mayoría de los alumnos (15 de ellos) tiene entre 3 y 5 hermanos.

Si se divide la frecuencia absoluta de cada clase (*las clases la constituyen los valores de la columna 1*) entre el total de observaciones (25 en este caso), se obtienen las frecuencias relativas (*columna 4*). Estas nos permiten conocer qué proporción de los alumnos tienen un número dado de hermanos.

Por ejemplo: $0,16$; que es la frecuencia relativa correspondiente a la tercera clase, indica que un 16% de los alumnos tienen 3 hermanos.

Las distribuciones de variables discretas pueden presentarse por medio de un gráfico de bastones o de un gráfico de barras verticales, utilizando las frecuencias absolutas o las relativas. Seguidamente se incluyen, para el ejemplo que estamos viendo, los dos tipos de gráficos empleando frecuencias absolutas.

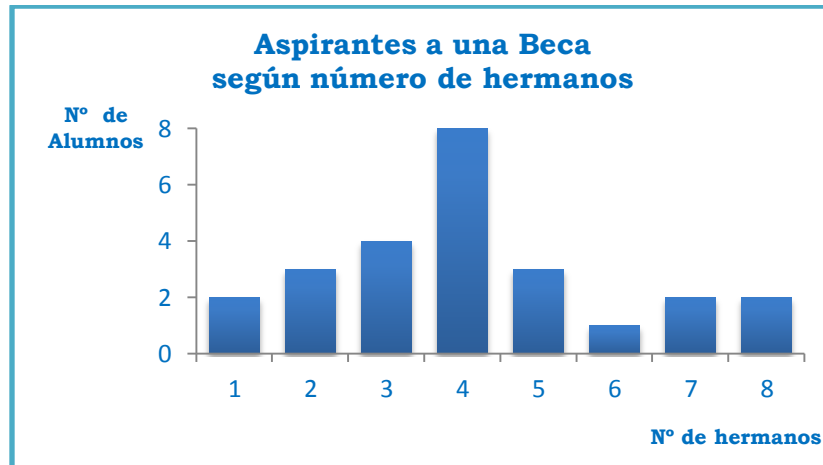


Figura 1: Gráfico de barras verticales para ilustrar la distribución de variable discreta del ejemplo 2.

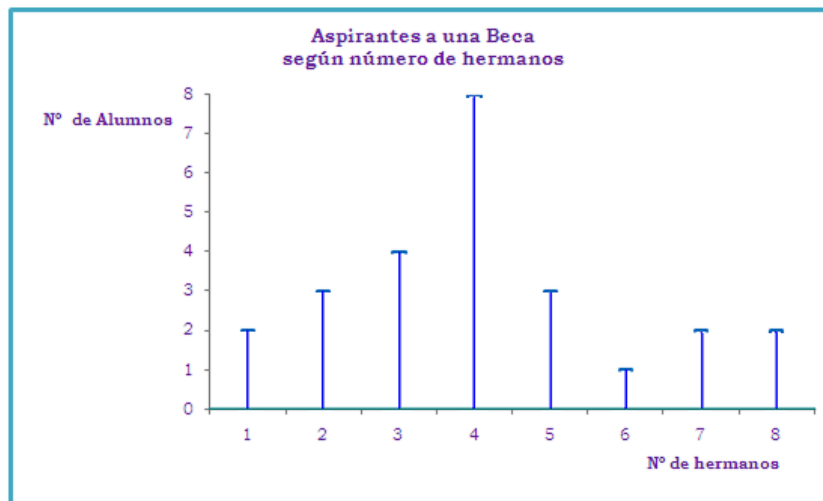


Figura 2: Gráfico de bastones para ilustrar la distribución de variable discreta del ejemplo 2.

Para ciertos propósitos resulta útil acumular las frecuencias absolutas (o las relativas). Esto puede hacerse sumándolas hacia abajo, para obtener una distribución acumulada “menos de” o sumándolas hacia arriba, para obtener una distribución acumulada “más de” (ver el siguiente cuadro).

Tabla 3: Distribuciones de frecuencias absolutas, relativas, “menos de” y “más de” para el ejemplo 2.

Nº HERMA NOS	FREC. ABS. (f_i)	FREC. REL. (f_r)	FRECUENCIA ACUMULADA “MENOS DE” (f_a)		FRECUENCIA ACUMULADA “MÁS DE” (f_a)	
			ABSOLUTA	RELATIVA	ABSOLUTA	RELATIVA
1	2	0,08	2	0,08	25	1,00
2	3	0,12	5	0,20	23	0,92
3	4	0,16	9	0,36	20	0,80
4	8	0,32	17	0,68	16	0,64
5	3	0,12	20	0,80	8	0,32
6	1	0,04	21	0,84	5	0,20
7	2	0,08	23	0,92	4	0,16
8	2	0,08	25	1,00	2	0,08
	25	1,00				

Por ejemplo: si se suman las frecuencias absolutas de las clases primera, segunda, tercera y cuarta, se obtiene $2 + 3 + 4 + 8 = 17$, valor que representa la frecuencia acumulada “*menos de*” de la cuarta clase. Si se realiza la misma operación con las frecuencias relativas se obtiene $0,08 + 0,12 + 0,16 + 0,32 = 0,68$.

Lo cual quiere decir que 17 estudiantes tienen 4 hermanos o menos (*esto lo sabemos porque se sumaron las categorías 1, 2, 3 y 4*); aunque también sería correcto decir que 17 estudiantes tienen “*menos de*” 5 hermanos. La interpretación del 0,68 es similar: 68% de los estudiantes tienen 4 *hermanos o menos*.

Si la suma se realiza de abajo hacia arriba, partiendo de la última clase se obtiene una frecuencia acumulada “*más de*”. Los números 5 y 0,20 de la sexta clase son ejemplo de ellas. Estos valores se obtuvieron así: el número 5 se obtuvo sumando las frecuencias absolutas de la octava, séptima y sexta clase $2 + 2 + 1 = 5$ y el número 0,20 sumando $0,08 + 0,08 + 0,04 = 0,20$.

La interpretación es la siguiente: el número 5 nos indica que 5 estudiantes tienen 6 hermanos o más (*aunque también sería correcto decir que 5 estudiantes tienen más de 5 hermanos*). El 0,20 se interpreta diciendo que el 20% de los estudiantes tienen 6 *hermanos o más*.

Ambos tipos de frecuencias acumuladas, la “*menos de*” y la “*más de*” son útiles; sin embargo, la más empleada en la práctica es la “*menos de*”.

La representación gráfica de las frecuencias acumuladas discretas (*absolutas y relativas*) puede hacerse recurriendo a gráficos de bastones o barras verticales. Cuando las clases son numerosas, sin embargo, algunas personas prefieren usar la curva simple para representar las acumuladas discretas.

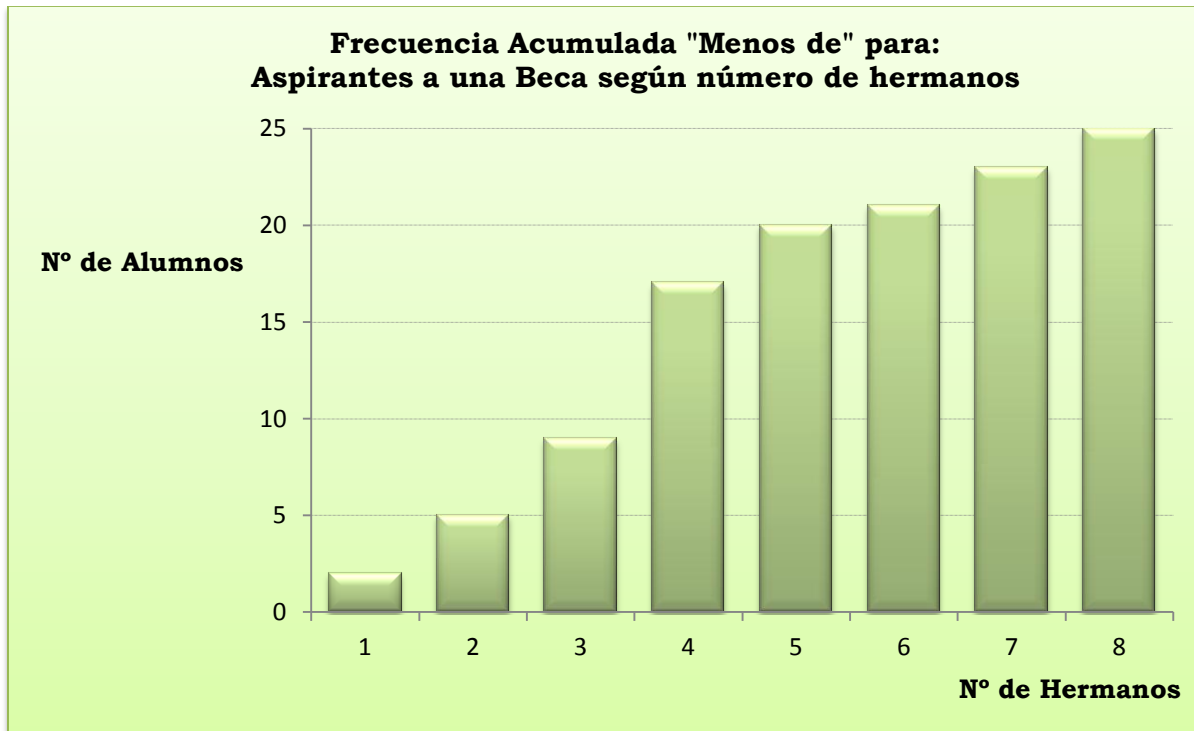


Figura 3: Gráfico de barras para ilustrar la frecuencia "menos de" (de variable discreta) del ejemplo 2.

Por lo general, las distribuciones de frecuencias para variables discretas son de muy fácil construcción; sin embargo, cuando se trata de variables continuas aparecen numerosas complicaciones y dificultades, ya que hay que considerar detalles o aspectos que deben ser tratados con especial atención y cuidado, tales como el número de clases que debe hacerse y su amplitud, pautas que deben seguirse al fijar los límites, etc.

Referencias

- ✚ GÓMEZ, MIGUEL. Estadística Descriptiva, Oficina de publicaciones de la Universidad de Costa Rica, 1977.
- ✚ QUINTANA, CARLOS. Estadística Elemental, Editorial Máster Libro S.A.1992.
- ✚ MOYA, LIGIA, Introducción a la Estadística de la Salud, Talleres de Litografía Punto Uno, 1994.