

16. Variables numéricas

□ 16.1. Histogramas y distribuciones de frecuencias

La **distribución** de una variable nos dice **cuáles son los valores** que puede tomar y su **frecuencia**, es decir, cuántas veces ocurre cada uno de los valores.

Como hemos visto, las tablas de frecuencias y los gráficos (circulares, de barras) permiten conocer la distribución (ya sea en una población o en una muestra) de los valores de una variable categórica. La distribución de los valores de la variable dentro de las diferentes categorías se puede expresar en cantidades, en proporciones o en porcentajes.

Para representar gráficamente la distribución de los datos correspondientes a una **variable numérica** (discreta o continua) también se utilizan tablas de frecuencias y un gráfico similar al gráfico de barras: el histograma.

Un **histograma** representa, en el eje horizontal, **los valores de una variable numérica** divididos en **intervalos de clase**. En forma similar a los gráficos de barras, tiene una barra sobre cada intervalo cuya **altura indica la cantidad** (frecuencia) o **proporción** (frecuencia relativa) de datos. No se deja espacio entre las barras ó rectángulos.

Cuando los valores posibles de la **variable numérica** son pocos, la altura de cada rectángulo del histograma muestra directamente la cantidad o proporción de veces que **cada uno de los valores** ocurrió. Cuando son muchos, es necesario agruparlos definiendo previamente los intervalos.

16.1.1. Variables discretas

Una variable numérica es **discreta** cuando únicamente puede tomar valores dentro de una sucesión determinada de números. La cantidad de hermanos por alumno de una escuela es una variable discreta: puede tomar los valores 0, 1, 2, 3, 4, pero nunca valores como 2,50; 7,2; 0,30.

Veremos primero un ejemplo de una variable numérica discreta (cantidad de hijos) con **pocos valores posibles. No es necesario agruparlos.**

Ejemplo 16.1. Supongamos que se entrevistan 1.000 familias de la Ciudad de Buenos Aires, para saber cuántos hijos tiene cada familia. Nuestros datos son de la forma 0, 0, 3, 1, 1, 1, 2, 2, 2, 3, 1, 1, 2, 0, 0, 0, 2, 1, 8, 1, 1, 2, 3, 0, 0, 0...

Cada número es la cantidad de hijos de cada una de las familias entrevistadas. Es necesario resumir la información: 250 familias no tienen hijos, 200 tienen 1 hijo, 300 tienen 2 hijos, 160 tienen 3 hijos, 50 tienen 4 hijos, 20 tienen 5 hijos, 10 tienen 6 hijos, 7 tienen 7 hijos, 2 familias tienen 8 hijos y una familia tiene 9 hijos. Podemos presentar el resumen mediante la siguiente tabla de frecuencias:

Tendremos una visualización más rápida de los datos si los representamos mediante un histograma.

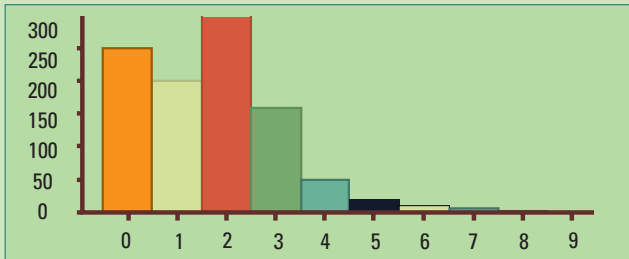


Figura 16.1. Histograma de la cantidad de hijos por familia, expresado en frecuencias.

Cantidad de hijos	Frecuencia
0	250
1	200
2	300
3	160
4	50
5	20
6	10
7	7
8	2
9	1
Total	1.000

La mayor cantidad de familias tienen 2 hijos, le siguen las familias sin hijos y después las de un sólo hijo.

Un histograma representa la **distribución de una variable numérica en una población o en una muestra**. Los intervalos de clase de una variable discreta están centrados en sus valores posibles y tienen la misma longitud.

En el ejemplo 16.1 los datos corresponden a una muestra de 1.000 familias de la Ciudad de Buenos Aires.

¿Cuál es la variable numérica y cuál es la población? ¿Cuáles son los valores posibles de esa variable numérica en la población? ¿Cuál es el tamaño de la muestra?:

- Variable numérica discreta: cantidad de hijos por familia.
- Valores posibles: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.
- Población: todas las familias de la Ciudad de Buenos Aires, en un año fijo.
- Tamaño de la muestra: 1.000

Si la muestra es representativa de las familias de la Ciudad de Buenos Aires en ese momento, podremos considerar al histograma, una estimación de la distribución de la variable cantidad de hijos por familia en la población. ¡Un verdadero trabalenguas!

Cuando interesa comparar la frecuencia entre categorías, como ocurre con los diagramas de barras, puede ser más interesante que el eje vertical esté expresado en frecuencias

relativas (es decir proporciones). Por ejemplo, si queremos estudiar el comportamiento social respecto a la cantidad de hijos, saber que el 75% de las familias tienen como máximo dos hijos es más informativo que saber que son 750.

Cantidad de hijos	Frecuencia	Frecuencia relativa	Porcentaje
0	250	$250/1.000 = 0,250$	25,0
1	200	$200/1.000 = 0,200$	20,0
2	300	$300/1.000 = 0,300$	30,0
3	160	$160/1.000 = 0,160$	16,0
4	50	$50/1.000 = 0,050$	5,0
5	20	$20/1.000 = 0,020$	2,0
6	10	$10/1.000 = 0,010$	1,0
7	7	$7/1.000 = 0,007$	0,7
8	2	$2/1.000 = 0,002$	0,2
9	1	$1/1.000 = 0,001$	0,1
Total	1000	1 1	100,0

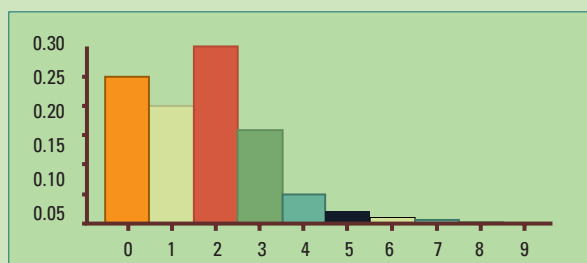


Figura 16.2. Histograma de la cantidad de hijos por familia, expresado en frecuencias relativas.

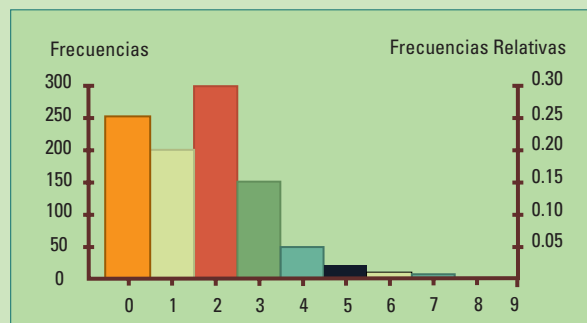


Figura 16.3. Histograma de la cantidad de hijos por familia, con dos escalas: Frecuencias y frecuencias relativas.

Observación. Los histogramas de frecuencias y de frecuencias relativas tienen siempre la misma forma, tal como se puede apreciar en las figuras 16.1 y 16.2. Cambian únicamente las escalas verticales. Algunas veces se presentan ambas en el mismo gráfico.

El ejemplo 16.1 (cantidad de hijos por familia) es hipotético. Como es difícil definir “familia”, resulta más realista considerar la cantidad de hijos por mujer, como veremos en el siguiente ejemplo con datos reales.

Ejemplo 16.2. Se trata de la cantidad de hijos de mujeres con edades entre 30 y 34 años en el año 1991 en la Ciudad de Buenos Aires (tabla 16.1); 25.729 mujeres no tienen hijos (24,5%), 19.573 mujeres tienen un solo hijo (18,6%), 33.060 mujeres tienen 2 hijos (31,4%), etc.

El ejemplo 16.1 (cantidad de hijos por familia) es hipotético. Como es difícil definir “familia”, resulta más realista considerar la cantidad de hijos por mujer, como veremos en el siguiente ejemplo con datos reales.

Ejemplo 16.2. Se trata de la cantidad de hijos de mujeres con edades entre 30 y 34 años en el año 1991 en la Ciudad de Buenos Aires (tabla 16.1); 25.729 mujeres no tienen hijos (24,5%), 19.573 mujeres tienen un solo hijo (18,6%), 33.060 mujeres tienen 2 hijos (31,4%), etc.

CANTIDAD DE HIJOS DE MUJERES, CON EDADES DESDE 30 A 34 AÑOS DE LA CIUDAD DE BUENOS AIRES. AÑO 1991. TABLA 16.1

Cantidad de hijos	Frecuencia	Frecuencia relativa	Porcentaje
0	25.729	$25.729/105.210 = 0,245$	24,5
1	19.573	$19.573/105.210 = 0,186$	18,6
2	33.060	$33.060/105.210 = 0,314$	31,4
3	18.020	$18.020/105.210 = 0,171$	17,1
4	5.467	$5.467/105.210 = 0,052$	5,2
5	1.867	$1.867/105.210 = 0,018$	1,8
6	813	$813/105.210 = 0,008$	0,8
7	380	$380/105.210 = 0,004$	0,4
8	216	$216/105.210 = 0,002$	0,2
9	85	$85/105.210 = 0,001$	0,1
Total	105.210	1	100,0

Fuente: Dirección General de Estadística y Censos (G.C.B.A.) sobre la base de datos del Censo Nacional de Población y Vivienda, 1991 - Serie C.

Un histograma de los datos de la tabla 16.1 nos permite visualizar más rápidamente su distribución.

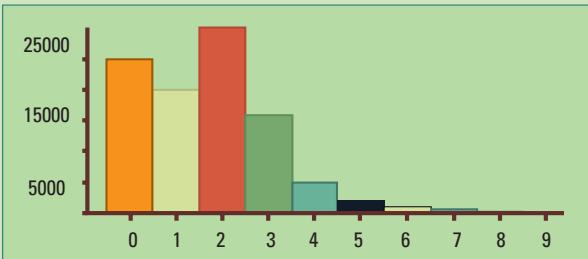


Figura 16.4. Datos reales. Ciudad de Buenos Aires año 1991. Histograma de la cantidad de hijos por mujer con edades entre 30 y 34 años.

La frecuencia (escala vertical del histograma, figura 16.4) es la cantidad de mujeres con edades entre 30 y 34 años en el año 1991, con 0,1, 2, ..., hasta 9 hijos, respectivamente en cada intervalo. Se destaca el rectángulo centrado en 2, porque tiene la mayor altura; 2 es la cantidad más frecuente de hijos en la Ciudad de Buenos Aires.

La distribución, es muy parecida a la del ejemplo hipotético; ambos histogramas tienen casi la misma forma pero las frecuencias, frecuencias relativas y porcentajes ya no son números redondos.

¿Cuál es la variable numérica y cuál es la población? ¿Cuáles son los valores posibles de esa variable numérica en la población?:

- Variable numérica discreta: cantidad de hijos por mujer
- Valores posibles: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 (no es posible tener 2,75 hijos).
- Población: todas las mujeres de la Ciudad de Buenos Aires entre 30 y 34 años en el 1991.

¿Puede haber mujeres con más de 9 hijos? Efectivamente, puede haber mujeres con 10 ó más hijos. En la ciudad de Buenos Aires sólo se incluye una categoría de 10 o más, porque son pocas. Para poder comparar las categorías mediante un histograma es necesario que tengan el mismo tamaño; es decir, que correspondan a la misma cantidad de valores posibles de la variable. Por esta razón no se incluyó en el histograma la categoría 10 ó más, correspondiente a los valores 10, 11, 12, 13, 14, etc.

16.1.2. Variables continuas

Una variable numérica es continua cuando, dados dos valores posibles de la variable, ésta siempre puede tomar cualquier valor intermedio. El peso de una persona es una variable numérica continua, puede tomar valores como 48 kg ó 49 kg y también, 48,5 kg 48,52 kg etc.

Podemos preguntarnos: ¿cambió la edad a la cual las mujeres tienen hijos? Veamos un ejemplo real para intentar responder esta pregunta. Como la variable edad tiene muchísimos valores posibles, para construir un histograma, los agruparemos en intervalos.

Ejemplo 16.3 Comparemos como se distribuye la edad de las mujeres en el momento del nacimiento de un hijo, en los años 2001, 2003, 2006, utilizando la información del Ministerio de Salud.

NACIMIENTOS EN LA REPÚBLICA ARGENTINA SEGÚN EDAD DE LA MADRE. TABLA 16.2

Año	2001	2003	2006	2001	2003	2006
Grupo de edad	Cantidad			Porcentaje		
[10-15)	3.022	2.763	2.766	0,44	0,40	0,40
[15-20)	97.060	92.461	103.885	14,20	13,25	14,92





Año	2001	2003	2006	2001	2003	2006
Grupo de edad	Cantidad			Porcentaje		
[20-25)	188.415	184.155	174.342	27,57	26,39	25,03
[25-30)	170.748	179.107	176.931	24,98	25,66	25,40
[30-35)	128.521	137.359	139.003	18,80	19,68	19,96
[35-40)	68.162	71.497	73.177	9,97	10,24	10,51
[40-45)	19.658	20.674	19.866	2,88	2,96	2,85
[45-50)	1.417	1.438	1.405	0,21	0,21	0,20
[50-55)	98	92	83	0,01	0,01	0,01
Sin información	6.394	8.406	4.993	0,94	1,20	0,72
Total	683.495	697.952	696.451	100,00	100,00	100,00

Fuente: Estadísticas Vitales. Ministerio de Salud. 2001, 2003, 2006. ISSN 1668-9054.

¿Cómo se interpretan los grupos de edad?

El grupo [10-15) corresponde a las edades entre 10 y 15 años

El grupo [15-20) corresponde a las edades entre 15 y 20 años

El grupo [20-25) corresponde a las edades entre 20 y 25 años

El grupo [25-30) corresponde a las edades entre 25 y 30 años

El grupo [30-35) corresponde a las edades entre 30 y 35 años

.....

Una edad de 15 años se cuenta en el grupo [15-20) y no en el [10-15)

Una edad de 20 años se cuenta en el grupo [20-25) y no en el [15-20)

.....

El intervalo [15-20) es un intervalo cerrado en 15 (se incluye el valor 15 en el intervalo) y abierto en 20 (no se incluye el valor 20 en el intervalo).

¿Cuál es la variable numérica y cuál es la población?:

En general, el intervalo **[a-b)**, donde **a** y **b** son números reales cualesquiera con **a** menor que **b**, es un **intervalo cerrado en a** (incluye el valor **a**) y **abierto en b** (no incluye el valor **b**)

- Variable numérica continua: edad de la madre en el momento del parto. Es posible tener una edad decimal de 18,75 años (18 años y 9 meses).
- Valores posibles: desde 10 hasta 54 años.
- Población: se consideran en este ejemplo tres poblaciones:
- Todos los niños nacidos en el año 2006.
- Todos los niños nacidos en el año 2003.
- Todos los niños nacidos en el año 2001.

Los histogramas de la figura 16.5 permiten comparar cómo se distribuyen las edades de las madres de la República Argentina en la población de los niños nacidos en el año 2006, 2003 y 2001 respectivamente.

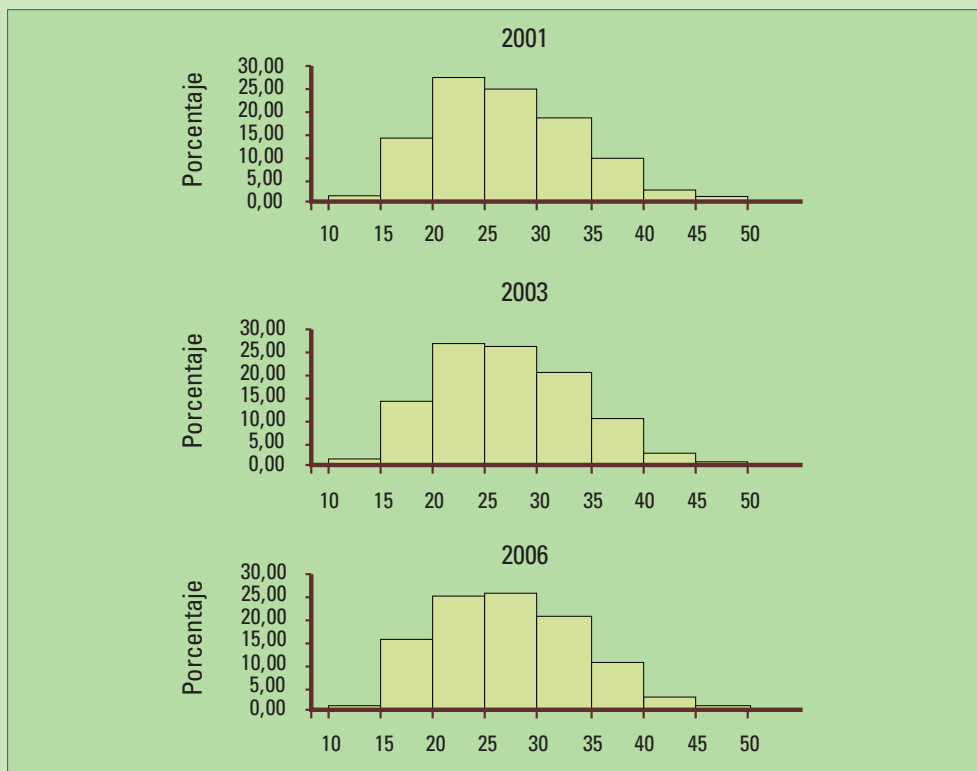
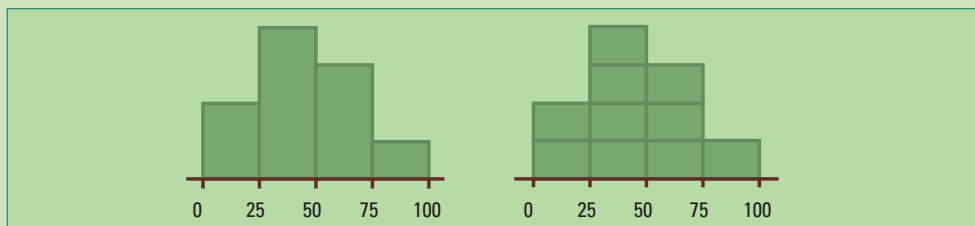


Figura 16.5. Edad de la madre en momento del parto para los años 2001, 2003, 2006 en la Ciudad Autónoma de Buenos Aires.

Los 3 histogramas de la figura 16.5 tienen formas similares, esto indicaría que la respuesta a la pregunta planteada es no. No cambiaron las edades en las cuales las mujeres tienen hijos en la República Argentina entre los años 2001, 2003 y 2006. Sin embargo, si observamos con más detalle vemos un porcentaje mayor en el año 2001 de nacimientos provenientes de madres con edades en el intervalo [20-25) años. En el 2003 esa diferencia entre los intervalos [20-25) y [25-30) se hace casi imperceptible y en el 2006 comienza ya el [25-30) tiene un porcentaje de 25,40 % un poco mayor que el del [20-25) con 25,03%. Mirando la tabla 16.2 (pág. 100) podemos ver además, porcentajes crecientes desde el 2001 al 2006 en los grupos de edades [30-35) y [35-40) desde el 2001 al 2006. Estas tendencias favorecen la idea que las mujeres tienen sus hijos a edades cada vez más tardías aunque se mantiene alto, cercano al 15%, el porcentaje de madres adolescentes. Esto es una preocupación de las autoridades sanitarias. La incidencia de prematuros, bajo peso al nacer y de parto instrumentado, es mayor entre las madres adolescentes que en madres con edades entre 20 y 30 años.

En un histograma puede faltar el eje vertical.

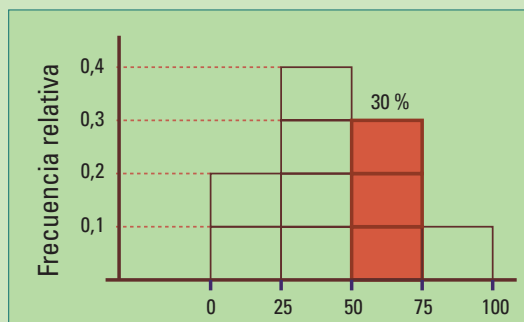
Ejemplo 16.4. Al siguiente histograma le falta el eje vertical. ¿Qué información puede proveer?



Sin el eje vertical no se pueden hallar las frecuencias absolutas, pero, sí es posible determinar las frecuencias relativas de cada uno de los intervalos. Debemos ver qué proporción del área total del histograma se encuentra por encima de cada intervalo. Dividimos la superficie del histograma en 10 rectángulos iguales de manera que cada porción es $1/10$ de esa superficie, es decir el 10%.

Hay 2 rectángulos sobre el intervalo 0-25, tienen el 20% del área; 4 rectángulos sobre 25-50, 40% del área; **3 rectángulos sobre 50-75, 30%** y 10% está sobre 75-100.

Generalmente, no es tan fácil dividir a los histogramas en 10 partes iguales, sin embargo siempre las frecuencias relativas se corresponden con áreas relativas.



□ 16.2. Construcción de tablas de frecuencias

En los ejemplos anteriores los datos ya estaban agrupados o los histogramas estaban contruidos. Vimos tablas con distribuciones de frecuencias para variables numéricas discretas (ejemplo 16.1 y 16.2) y para una variable numérica continua (ejemplo 16.3).

En las siguientes secciones veremos cómo se agrupan los datos numéricos y se construyen las tablas de frecuencias para obtener los histogramas. Trataremos en forma separada a los datos de variables discretas y continuas.

16.2.1. Variable discreta

Paso 1. Se ordenan los valores posibles de la variable.

Paso 2. Se cuenta cuántas veces aparece un dato con cada valor posible. Esto nos da la frecuencia.

Paso 3. Se divide cada frecuencia por el total de datos, obteniendo así la frecuencia relativa.

Ahora su turno: Registre cuántos hermanos tienen cada uno de los alumnos de su división y obtenga una tabla de frecuencias y de frecuencias relativas. ¿Cuál es la variable? ¿Cuáles son sus valores posibles? A partir de la tabla construya el histograma correspondiente. ¿Cuál es la población en estudio?

16.2.2. Variable continua

Paso 1. Se ordenan los datos.

Paso 2. Se definen intervalos de clase con igual longitud, cubriendo el rango de los valores observados.

Paso 3. Se cuentan cuantos datos pertenecen a cada uno de los intervalos. Esto indica la frecuencia.

Paso 4. Se divide cada frecuencia por el total de datos, obteniendo así la frecuencia relativa.

En el ejemplo siguiente veremos cómo construir la tabla de frecuencias para datos de una variable numérica continua.

Ejemplo 16.5. Los datos siguientes corresponden al peso (en kg) de 52 alumnos y 49 alumnas de 3 divisiones de 4to. año.

- **Varones:** 67 57 64 73 65 69 67 66 67 69 63 65 66 53 58 64 69 67 63 71 69 62 59 61 72 68 57 55 79 59 66 58 72 67 71 67 65 61 63 69 74 64 66 70 63 51 79 68 67 66 85 81
- **Mujeres:** 46 52 52 52 51 43 48 44 55 43 50 57 52 54 51 54 48 48 62 52 50 52 45 54 47 50 50 51 60 56 51 52 54 42 54 48 50 56 50 48 52 55 54 58 46 37 38 68 70

¿Cuál es la variable? Peso

¿Es una variable numérica continua o discreta? El peso es una variable numérica continua.

¿Cuál es la población?

Si nos interesa describir el peso de los/as alumnos/as de esas 3 divisiones de 4to. año, la población está formada por todos/as los alumnos/as de esas 3 divisiones.

¿Qué podemos decir de la distribución de los pesos mirando estos datos?

Para comenzar construiremos un diagrama de puntos, donde cada punto corresponde a un alumno de ese peso. Los valores repetidos se ponen uno encima del otro, a iguales distancias. ¿Se puede ver algo raro? Hay espacios vacíos y se distinguen 2 picos.

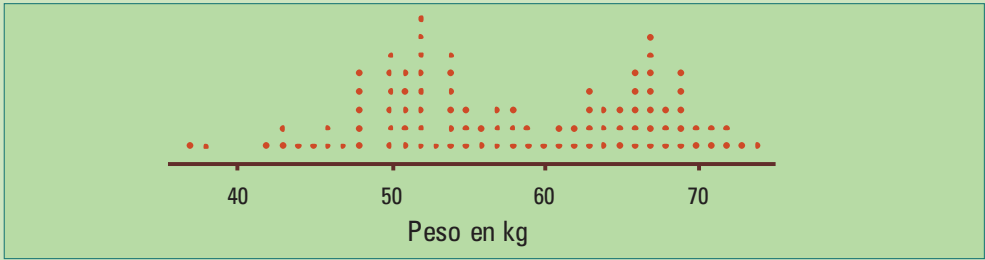


Figura 16.6. Diagrama de puntos de los pesos de varones y mujeres de 4to. año.

Luego, construiremos una tabla de frecuencias, para eso se dividimos la recta numérica en intervalos de clase y contamos cuántos pesos caen dentro de esos intervalos. La frecuencia relativa es la proporción de pesos dentro de cada intervalo.

**FRECUENCIAS DE LOS PESOS (EN kg)
DE LOS ALUMNOS Y ALUMNAS DE
4TO. AÑO.** TABLA 16.3

Intervalo de Clase	Frecuencia	Frecuencia relativa
[30 - 45)	6	
[45 - 60)	48	
[60 - 75)	43	
[75 - 90)	4	
		COMPLETAR
Total	101	1

El intervalo [30-45) es un intervalo cerrado en 30 (se incluye el valor 30 en el intervalo) y abierto en 45 (no se incluye el valor 45 en el intervalo).

El intervalo [45-60) es un intervalo cerrado en 45 (se incluye el valor 45 en el intervalo) y abierto en 60 (no se incluye el valor 60 en el intervalo).

El número al lado del corchete se incluye en el intervalo, el número al lado del paréntesis no.

Ahora su turno. Completar:

El intervalo [60-75) es un intervalo cerrado en y abierto en, porque

El intervalo [75 - 90) es un intervalo cerrado en y abierto en, porque

FRECUENCIAS DE LOS PESOS (EN kg) DE LOS ALUMNOS DE 4TO. AÑO. TABLA 16.4

Intervalo de Clase	Frecuencia	Frecuencia relativa
[30 - 35)	0	
[35 - 40)	2	
[40 - 45)	4	
[45 - 50)	9	
[50 - 55)	26	
[55 - 60)	13	
[60 - 65)	12	
[65 - 70)	23	
[70 - 75)	8	
[75 - 80)	2	
[80 - 85)	1	
[85 - 90)	1	
		COMPLETAR
Total	101	1



Figura 16.7.

Ahora, se debe construir el histograma. Éste (figura 16.7) no parece demasiado interesante. La mayoría de los pesos se encuentran entre los 45kg y los 75 kg, entonces podemos subdividir los intervalos de clase en tres partes iguales y obtenemos una nueva tabla de frecuencias (tabla 16.4).

El primer intervalo de clase [30-35) no tiene datos, por lo tanto ningún/a alumno/a tiene su peso dentro de ese intervalo. ¿Qué significan el corchete y el paréntesis?

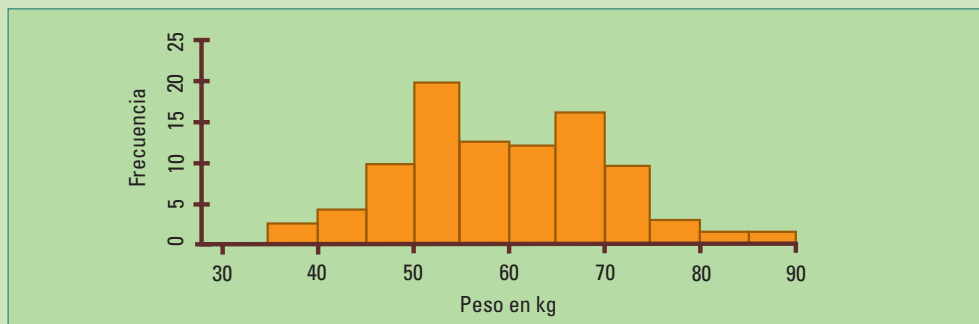


Figura16.8. Histograma los pesos de varones y mujeres de 4to. año.

Ahora el histograma (figura 16.8), de manera similar al diagrama de puntos (figura 16.6), nos muestra una información más interesante de la distribución de los pesos. Ambos sugieren la presencia de dos grupos aunque no se vean totalmente separados. En este ejemplo, conocemos los dos grupos mezclados, varones y mujeres. En el histograma se puede apreciar además, el carácter continuo de la variable peso.

No hay una regla para obtener la cantidad más conveniente de intervalos de clase, pero daremos unas ideas al respecto:

- Utilice intervalos de igual longitud centrados en valores redondos, si es posible, enteros.
- Si tiene pocos datos utilice una pequeña cantidad de intervalos.
- Para conjuntos de datos más grandes utilice más cantidad de intervalos.
- Una cantidad adecuada suele ser entre 6 y 12 intervalos.

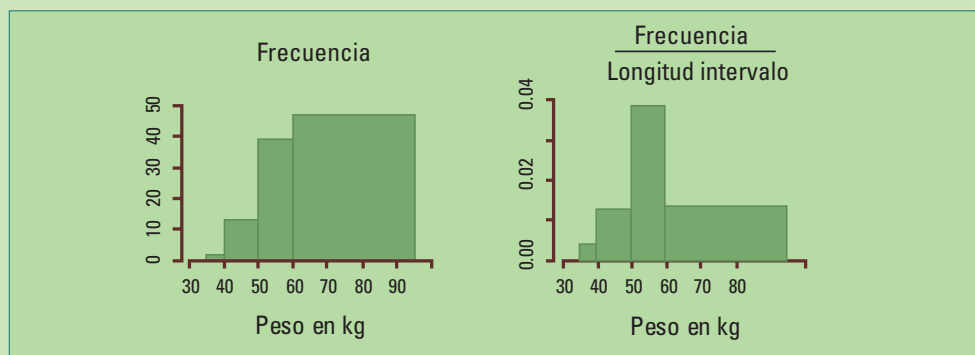
16.2.2.1. Un detalle extra

¿Pueden los **intervalos de clase** de un histograma tener **longitudes diferentes**?

Pueden, pero su construcción se complica.

En ese caso, para la altura del rectángulo de cada clase es necesario utilizar la frecuencia o la frecuencia relativa **dividida** por **la longitud** de dicho **intervalo de clase** (llamada **escala densidad**), de lo contrario, aumentar la longitud implicaría aumentar la altura, y disminuir su longitud resultaría en reducir la altura, **distorsionando** artificialmente la forma del histograma.

La figura siguiente muestra dos histogramas, en el de la izquierda la escala vertical es la frecuencia, y en el de la derecha, la frecuencia relativa dividida la longitud del intervalo de clase.



En el histograma de la izquierda, de **frecuencias absolutas** de los pesos de alumnas y alumnos de 4to. año, utilizando **intervalos de clase de distinta longitud**, **no representa** adecuadamente **la distribución de los datos** (ver figuras 16.7 y 16.8). Muestra más alumnos entre 60 y 90 kg que entre 30 y 60 kg. El de la derecha mejora la representación de la distribución de los datos.

Conclusión. Siempre que pueda **utilice intervalos de clase de la misma longitud**. Si no es posible elija la escala de densidad para el eje vertical.

□ 16.3. Diagrama tallo - hoja

Los histogramas son adecuados para conjuntos grandes de datos. Muestran su distribución pero se pierden los valores individuales. Para conjuntos con alrededor de 100 datos o menos, preferimos utilizar un diagrama tallo-hoja pues muestra no sólo la distribución de los datos sino también sus valores.

El estadístico John Tukey propuso en 1975, los diagramas tallo-hoja, una forma rápida para mostrar la distribución de datos correspondientes a variables numéricas, sin necesidad de obtener tablas de frecuencias, conservando todos los valores.

En estos diagramas las filas juegan el mismo papel de los rectángulos de clase en un histograma. Son como un **histograma girado 90°**. Cada fila está encabezada por un número, llamado **tallo**, a continuación se coloca una **línea vertical** y luego **las hojas**. Los valores de los tallos indican en forma compacta los intervalos de clase y tienen valores crecientes hacia abajo. Las hojas representan a los datos.

A continuación, construimos un diagrama tallo-hoja con los datos del ejemplo 16.5, el peso de alumnos y alumnas:

- **Varones:** 67 57 64 73 65 69 67 66 67 69 63 65 66 53 58 64 69 67 63 71 69 62 59 61 72 68 57 55 79 59 66 58 72 67 71 67 65 61 63 69 74 64 66 70 63 51 79 68 67 66 85 81
- **Mujeres:** 46 52 52 52 51 43 48 44 55 43 50 57 52 54 51 54 48 48 62 52 50 52 45 54 47 50 50 51 60 56 51 52 54 42 54 48 50 56 50 48 52 55 54 58 46 37 38 68 70

Intervalo	Tallo	Intervalo Tallo	Tallo
[30, 35)	3	[60, 65)	6
[35, 40)	3	[65, 70)	6
[40, 45)	4	[70, 75)	7
[45, 50)	4	[75, 80)	7
[50, 55)	5	[80, 85)	8
[55, 60)	5	[85, 90)	8

Elegimos los intervalos de clase y les asignamos su tallo

Los tallos están repetidos, aparecerán en el diagrama en dos filas consecutivas. En la fila superior van las hojas desde el cero al 4 y en la inferior las hojas desde el 5 al 9. Por ejemplo, el 5 de la fila superior representa al intervalo [50, 55] y allí se colocan las hojas (el segundo dígito) de todos los datos de ese intervalo y en la inferior se colocan las hojas de todos los datos del intervalo [55, 60].

El **tallo** es una columna de números correspondientes al primer dígito de los datos (dejamos el segundo dígito para las hojas)

Tallo los números crecen hacia abajo

3
3
4
4
5
5
6
6
7
7
8



En la segunda fila con tallo 5 se colocan 7 8 representando 57 kg 58 kg

Colocamos **el segundo dígito**, la hoja, en la fila adecuada

Tallo Hojas

3
3
4
4
5
5
6
6
7
7
8

3
78
43
759767956
3

Hemos colocado los pesos de los primeros quince varones 67 57 64 73 65 69 67 66 67 69 63 65 66 53 58

Ya hemos completado el diagrama con todos los datos

Tallo Hojas

3
3 78
4 3432
4 688857886
5 31222102414202400112440024
5 7897598576658
6 43432113320
6 7597679569798675968768
7 3122400
7 99
8 1
8 5

Finalmente ordenamos los valores de las hojas

Tallo Hojas

3
3 78
4 2334
4 566788888
5 0000001111122222223444444
5 5556677788899
6 011223333444
6 5556666677777788899999
7 00112234
7 99
8 1
8 5