



Lección 19: Histogramas

En las lecciones anteriores representamos gráficamente algunos conjuntos de valores, ya sea con gráficas de barras o con gráficas circulares. Sin embargo con las variables numéricas continuas, estas formas de representación no son útiles, debido al tipo de valores que se presentan.

Al hablar de variables numéricas continuas nos referimos a aquellas que pueden tener entre sus valores cualquier número racional, por ejemplo si midiéramos en metros la *estatura* de personas adultas podríamos encontrar valores como 1.58, 1.47, 1.60, 1.45, 1.72, 1.80, etc.; algo similar ocurre con el *peso*, en este caso los valores que podríamos encontrar en personas adultas al medir su *peso* en kilogramos serían como 60, 57.300, 50, 83.500, etc.

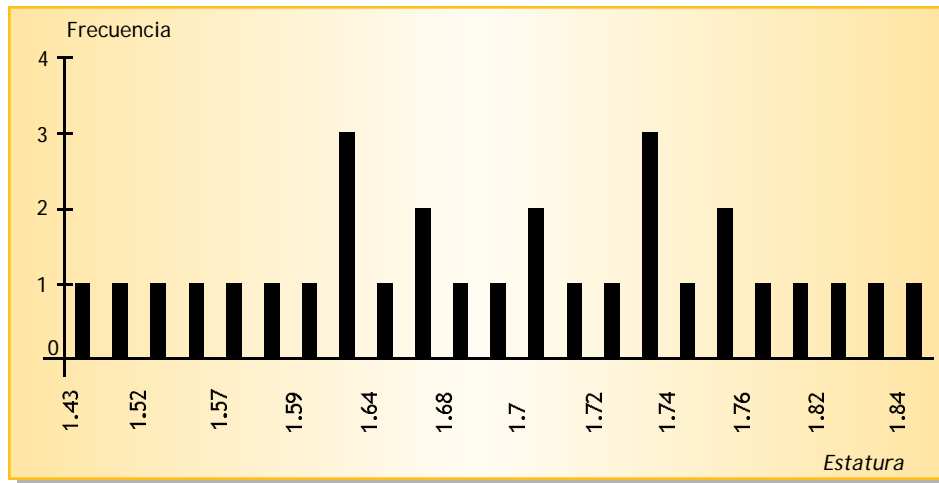
Cuando en la práctica tenemos que hacer mediciones de esta naturaleza nos encontramos que la gran mayoría de ellas son distintas y las mediciones iguales se repiten muy pocas veces, entonces la frecuencia con que aparece cada valor es 1 o un número muy pequeño. Además, como los valores varían mucho y debemos respetar la unidad de medida, necesitaríamos una gráfica muy larga y de barritas muy pequeñas que sería más difícil de leer que la misma colección de datos. Veamos un ejemplo:

En un grupo de 30 jóvenes de 18 a 20 años de una colonia del Distrito Federal, se midió la estatura y se obtuvieron los datos que se presentan a continuación, ordenados de menor a mayor:

1.43	1.58	1.64	1.70	1.73	1.76
1.49	1.59	1.65	1.70	1.73	1.77
1.52	1.63	1.65	1.71	1.74	1.82
1.55	1.63	1.68	1.72	1.75	1.83
1.57	1.63	1.69	1.73	1.75	1.84

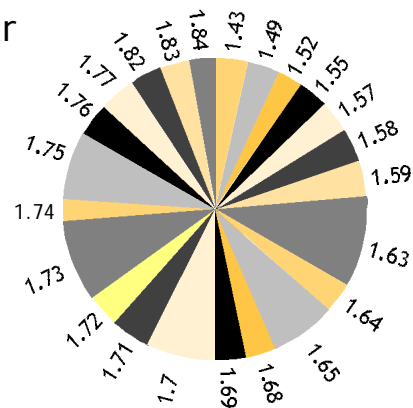


Al calcular las frecuencias podemos observar que los únicos valores que tienen frecuencias distintas de 1 son: 1.63 y 1.73, que tienen frecuencias iguales a 3, y 1.65, 1.70 y 1.75, que tienen frecuencias iguales a 2. La gráfica de barras correspondiente a estos datos es la que se muestra aquí.



Considerando que la finalidad de una gráfica es permitirnos visualizar, fácilmente, el conjunto de datos y percibir las relaciones entre valores y frecuencias, es claro que esta gráfica no cumple con su objetivo.

Aún puede ser más complicado tratar de leer una gráfica circular que contenga estos datos:



Para poder construir una gráfica que realmente aporte información, se agrupan los valores en intervalos, esto es consideramos juntos los valores que pertenecen a un mismo intervalo. Por ejemplo, si nos referimos a los valores que son mayores que 1.40 y menores o iguales 1.50, estamos hablando de los dos primeros valores de la lista.

Veamos como construir una tabla de frecuencias para valores agrupados en intervalos (que también se denominan clases). Primero tenemos que decidir cuáles serán los intervalos, es decir, dónde empieza y termina cada uno de ellos; aquí hay que tener en cuenta lo siguiente:

- Todos los intervalos deben tener la misma longitud, esto es las distancias entre los extremos de los intervalos deben ser las mismas en todos.
- Cada valor debe pertenecer sólo a un intervalo. Esto significa que ningún dato debe quedar fuera de la agrupación y que ningún dato puede pertenecer a dos intervalos distintos.

Para que nos quede cómodo nosotros vamos a agrupar los datos de diez en diez centímetros: de 1.40 a 1.50, de 1.50 a 1.60, 1.60 a 1.70, y así sucesivamente. Como las expresiones anteriores no nos dicen dónde ubicar algunos números, como por ejemplo el 1.70; recurrimos a los intervalos semiabiertos que usted estudió en la lección 3 de este libro. Vamos a usar intervalos abiertos por la izquierda y cerrados por la derecha, por ejemplo el intervalo $(1.40, 1.50]$; como ya sabemos en este intervalo están todos los números mayores que 1.40 y menores o iguales a 1.50. De los datos que nosotros tenemos los que pertenecen al intervalo anterior son 1.43 y 1.49; sabemos también que 1.70 está dentro del intervalo $(1.60, 1.70]$.

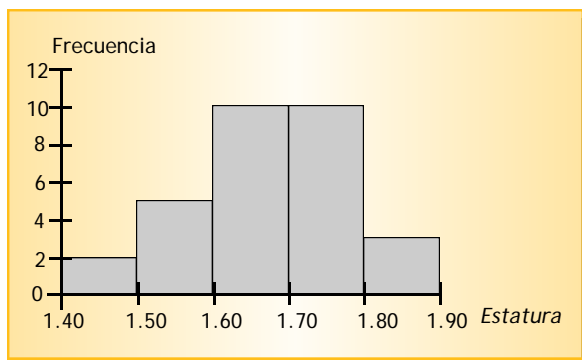
Para calcular la frecuencia de cada intervalo contamos cuántos datos de los que tenemos pertenecen a ese intervalo. A continuación reproducimos la lista original de datos, pero ahora ponemos en columnas distintas los datos que pertenecen a distintos intervalos:

1.43	1.52	1.63	1.71	1.82
1.49	1.55	1.63	1.72	1.83
1.52	1.57	1.63	1.73	1.84
1.55	1.58	1.64	1.73	
1.57	1.59	1.65	1.73	
		1.65	1.74	
		1.68	1.75	
		1.69	1.75	
		1.70	1.76	
		1.70	1.77	

Veamos entonces cómo queda nuestra tabla de frecuencias:

Intervalo o clase	Frecuencia	Frecuencia relativa
(1.40, 1.50]	2	0.07 = 7%
(1.50, 1.60]	5	0.17 = 17%
(1.60, 1.70]	10	0.33 = 33%
(1.70, 1.80]	10	0.33 = 33%
(1.80, 1.90]	3	0.10 = 10%
Totales	30	1 = 100%

A partir de la tabla de frecuencias construimos la siguiente gráfica:

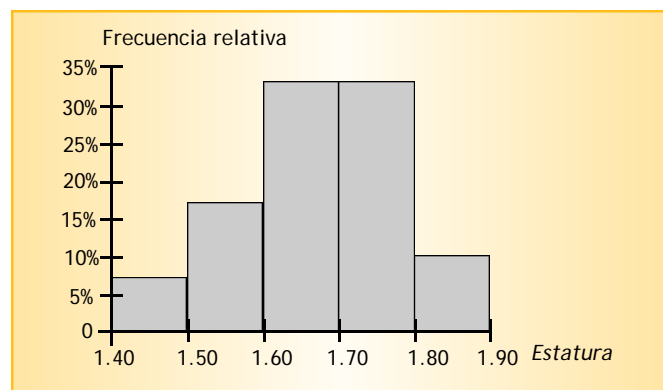


Una gráfica como ésta recibe el nombre de **histograma**. Puede observar que una vez establecidas las clases (o intervalos) el proceso de construcción es similar al de las gráficas de barras. La diferencia en el dibujo está

en la posición de las barras: en el histograma las barras están pegadas, cada una comienza donde acaba la anterior, mientras que en las que habíamos elaborado anteriormente las barras van separadas unas de otras.

Con una gráfica de esta naturaleza es claro que perdemos la información de sobre los valores individuales, ya no sabemos cuáles fueron los valores obtenidos al medir la estatura de cada joven, pero ganamos en información global. Por ejemplo, una simple ojeada a la gráfica basta para detectar que la gran mayoría de los jóvenes midieron entre 1.60 m y 1.80 m.

Así como se construyó el histograma de frecuencias, también puede construirse el histograma de frecuencias relativas.



Ejercicio 1

Considere los datos de la variable tiempo de la tabla de datos de la lección anterior y complete la siguiente tabla:

clase	Frecuencia	Frecuencia relativa
(60, 70]		
(70, 80]		
(80, 90]		
(90, 100]		
(100, 110]		
Totales		

Ejercicio 2

A partir de la tabla anterior, construya:

- Un histograma de frecuencias.
- Un histograma de frecuencias relativas.

Ejercicio 3

Se les preguntó a los obreros de una fábrica cuánto tiempo empleaban para trasladarse desde su domicilio al lugar de trabajo. Con los datos obtenidos se construyó la tabla que se muestra a continuación.

clase	Frecuencia	Frecuencia relativa
[45, 55)	4	0.03
[55, 65)	16	0.11
[65, 75)	36	0.24
[75, 85)	60	0.40
[85, 95)	31	0.21
[95, 105)	0	0
[105, 115)	3	0.02
Totales	150	1

Con los datos de la tabla construya:

- Un histograma de frecuencias.
- Un histograma de frecuencias relativas.

Ejercicio 4

Con base en la información de la tabla anterior conteste las siguientes preguntas:

- ¿Cuántos obreros fueron consultados?
- ¿Cuántos obreros emplean entre 65 y 75 minutos en trasladarse de su domicilio al lugar de trabajo?
- ¿Cuántos obreros emplean entre 55 y 75 minutos en trasladarse de su domicilio al lugar de trabajo?
- ¿Cuántos obreros emplean entre 95 y 105 minutos en trasladarse de su domicilio al lugar de trabajo?
- ¿Cuántos obreros emplean más de 85 minutos en trasladarse de su domicilio al lugar de trabajo?
- ¿Cuántos obreros emplean menos de 75 minutos en trasladarse de su domicilio al lugar de trabajo?

