

Universidad de San Carlos de Guatemala
Facultad de Odontología
Área Básica
Estadística Básica

Elaboración de Histogramas y Polígonos de Frecuencia.

Guatemala, abril 2010.

Introducción

Los objetivos principales de la estadística descriptiva son: la recolección, el ordenamiento y la presentación de la información.

La elaboración de distribuciones de frecuencias y gráficas es una parte importante de la presentación de la información y dentro de las gráficas la realización, manejo e interpretación del histograma y el polígono de frecuencias resulta muy importante.

Recuerde que el objetivo de este tema no es simplemente la construcción de gráficas, sino mas bien aprender algo acerca de los conjuntos de datos, es decir, comprender la naturaleza de sus distribuciones.

Histogramas.

Definición: Histograma es una gráfica de áreas en donde en la escala horizontal se representan los intervalos de clase de los datos y en la escala vertical se ubican las frecuencias o frecuencias relativas. El área contenida en cada una de las celdas corresponde a los valores de frecuencia, por lo que todas las celdas se elaboran del mismo ancho y se dibujan de manera adyacente.

Elaboración del Histograma.

Para poder elaborar un histograma necesitamos tener con anterioridad una distribución de frecuencia de los datos.

Por ejemplo:

Intervalos	Frec.
10 a 19	19
20 a 29	14
30 a 39	8
40 a 49	5
50 a 59	3
60 a 69	2
70 a 79	1
Totales	52

1o. Debemos determinar la unidad de medida utilizada en los intervalos de clase.

Para esto restamos el límite superior del primer intervalo con el límite inferior del segundo intervalo de clase.

El límite superior del primer intervalo es 19

El límite inferior del segundo intervalo es 20

Por lo tanto la unidad de medida (lo que incrementa entre un intervalo y el otro) es 1.

2o. Debemos calcular los límites reales de los intervalos.

Para esto aplicamos la fórmula para los límites reales en una medición con la diferencia de que disminuirémos la mitad de la unidad de medición al límite inferior del intervalo y le aumentamos la mitad de la unidad de medición al límite superior.

$$L.R. = X \pm (U.M./2)$$

En este caso nuestra X será el intervalo de clase y la unidad de medida es 1.

$$L.R. = (10 \text{ a } 19) \pm (1/2) \implies \begin{aligned} L.R. &= 10 - 0.5 = 9.5 \\ L.R. &= 19 + 0.5 = 29.5 \end{aligned}$$

para el segundo intervalo sería:

$$L.R. = (20 \text{ a } 29) \pm (1/2) \implies \begin{aligned} L.R. &= 20 - 0.5 = 19.5 \\ L.R. &= 29 + 0.5 = 29.5 \end{aligned}$$

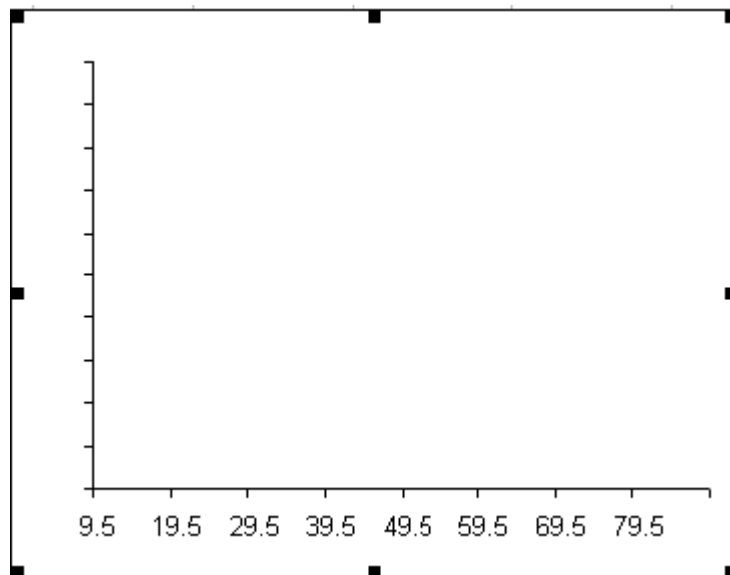
y así sucesivamente.

El resultado final será:

Intervalos	Frec.
9.5 a 19.5	19
19.5 a 29.5	14
29.5 a 39.5	8
39.5 a 49.5	5
49.5 a 59.5	3
59.5 a 69.5	2
69.5 a 79.5	1
Totales	52

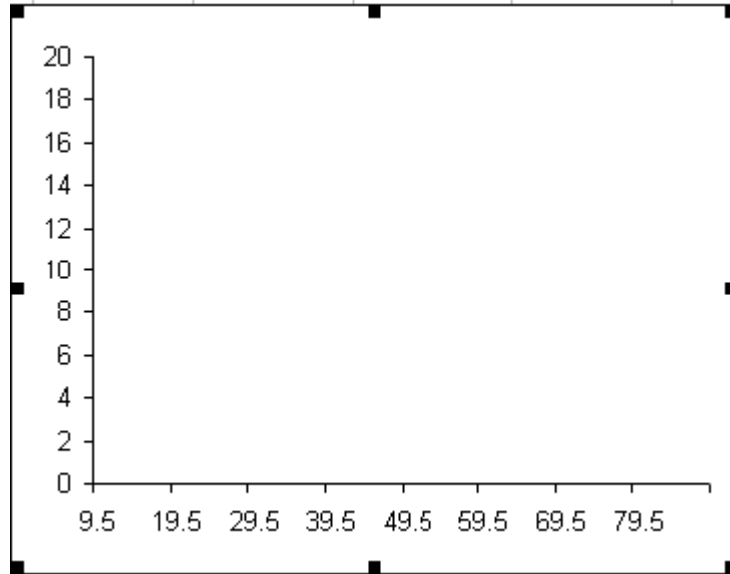
Luego colocamos los límites reales en el eje horizontal o eje X.

El eje vertical o eje Y, debe medir 3/4 partes de la longitud del eje de las X.

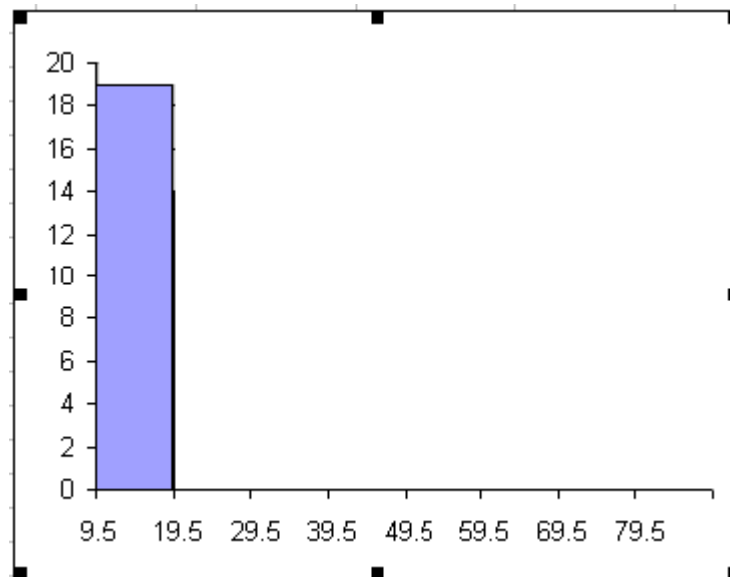


Luego ubicamos nuestra frecuencia más alta y la utilizamos guía para el valor máximo en el eje de las Y.

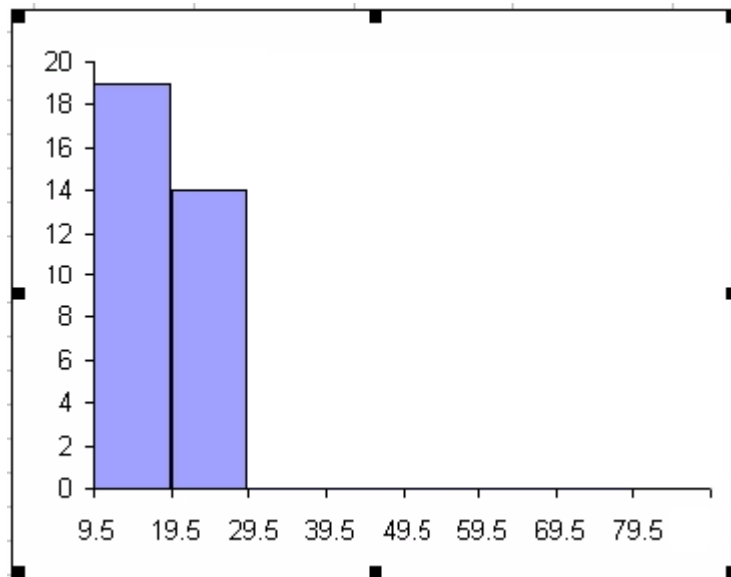
En este ejemplo nuestra frecuencia máxima es de 19, por lo que ubicamos el 20 como el valor más alto en el eje vertical. Luego ubicamos el resto de valores de manera proporcional.



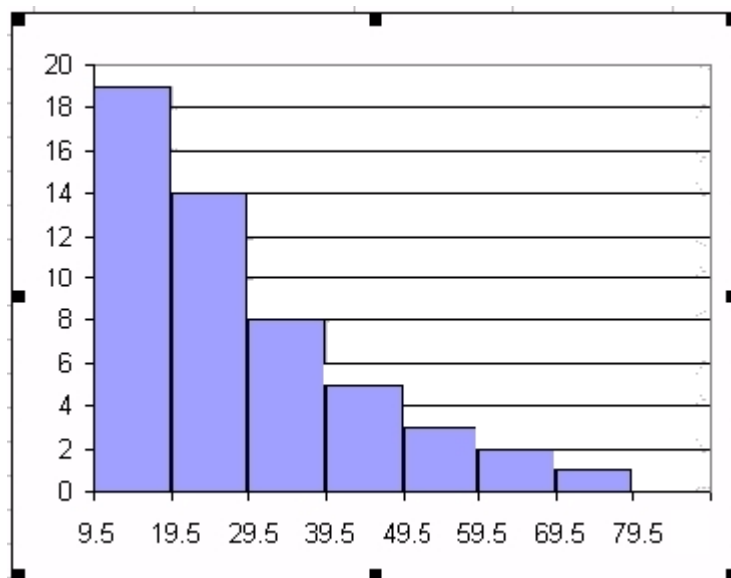
A continuación levantamos un rectángulo ubicado sobre el primer intervalo y hasta su frecuencia correspondiente.



Dibujamos el segundo rectángulo sobre el segundo intervalo, cuidando de no dejar espacio entre los rectángulos.



Al finalizar de ubicar nuestros rectángulos, tenemos terminado nuestro histograma.



Polígono de frecuencias.

El polígono de frecuencias es una gráfica de áreas que utiliza segmentos lineales conectados a puntos que se localizan directamente por encima de los valores de las marcas de clase. Las alturas de los puntos corresponden a las frecuencias de clase, en tanto que los segmentos lineales se extienden hacia la derecha y la izquierda a una marca de clase extra en cada sentido, de manera que la gráfica inicia y termina sobre el eje horizontal.

Elaboración del polígono de frecuencias.

Para poder elaborar un polígono de frecuencias necesitamos tener con anterioridad una distribución de frecuencia de los datos.

Por ejemplo:

Intervalos	Frec.
10 a 19	19
20 a 29	14
30 a 39	8
40 a 49	5
50 a 59	3
60 a 69	2
70 a 79	1
Totales	52

1°. Debemos obtener la marca de clase de cada intervalo. Esto es el punto medio de cada uno de los intervalos.

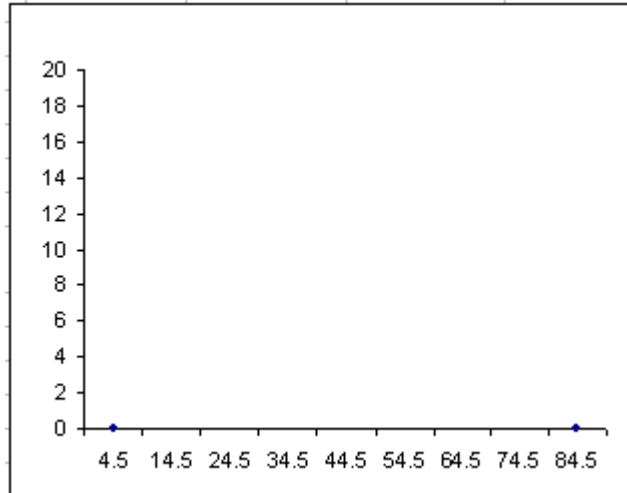
Lo obtenemos sumando los límites inferior y superior del primer intervalo y dividiendo el resultado entre de dos [$(Li + Ui) / 2$].

Así en nuestro ejemplo, la marca de clase del primer intervalo sería: $(10+19) / 2 = 14.5$

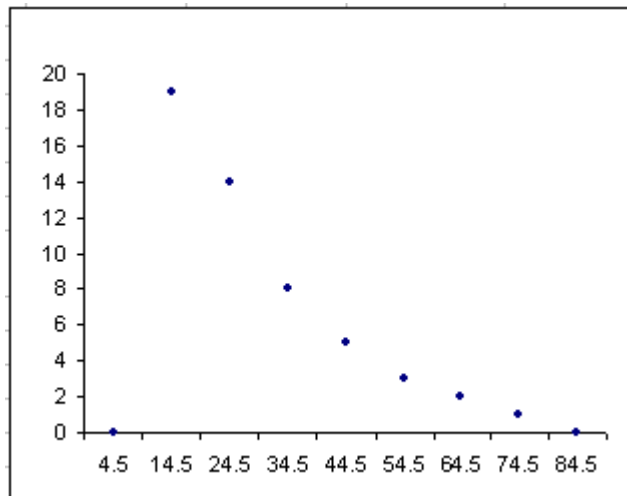
A este resultado le sumamos la amplitud del intervalo, que en este caso es de 10. Y continuamos sucesivamente hasta completar los intervalos. Agregamos una marca de clase anterior y una marca de clase posterior. El resultado será:

Intervalos	Marca de Clase	Frec.
	4.5	
10 a 19	14.5	19
20 a 29	24.5	14
30 a 39	34.5	8
40 a 49	44.5	5
50 a 59	54.5	3
60 a 69	64.5	2
70 a 79	74.5	1
	84.5	
Totales		52

2°. Ubicamos las marcas de clase en el eje horizontal, y en el eje vertical (que debe medir 3/4 partes de la longitud total del eje horizontal) ubicamos la frecuencia máxima.



A continuación ubicamos la posición a la que corresponde cada una de las frecuencias en sus respectivas marcas de clase.



Y por último unimos los puntos entre si, iniciando y terminando en el eje horizontal.

