

HISTOGRAMA

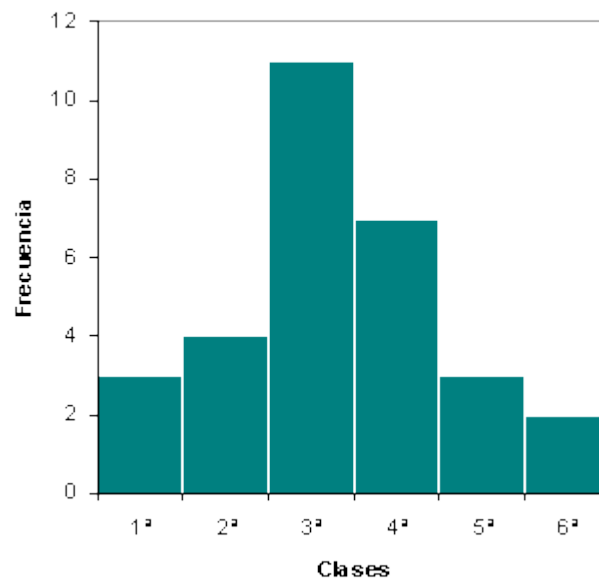
HISTOGRAMA

¿QUÉ ES EL HISTOGRAMA?

- El histograma es una herramienta útil para resumir y analizar datos.
- Por su naturaleza gráfica, puede ayudar a identificar e interpretar pautas que son difíciles de ver en simples tablas de números.
- Es una representación gráfica utilizada para visualizar y analizar la [frecuencia](#) con que una [variable](#) toma diferentes valores dentro de un conjunto de datos.

Frecuencia: Número de veces que se produce un suceso.

Variable: Característica cuantitativa de un suceso susceptible de ser medida y de ser representada por un valor numérico determinado, relacionado con una escala de medida.



¿CÓMO SE UTILIZA?

1. Recoger los datos X, que se van a analizar. Estos datos deben ser fiables y representativos de las condiciones típicas y normales del problema o del proceso, por lo que los datos no deben ser obsoletos o incompletos.

EJEMPLO

Una empresa debe fabricar tornillos que tienen como valor especificado de longitud $25 \pm 0,4$ mm. Para evaluar el número de piezas con errores de tolerancia se toman 30 muestras, tal y como se muestra en la [tabla](#).

Muestra	Longitud (mm)	Muestra	Longitud (mm)	Muestra	Longitud (mm)
1	25,2	11	25,3	21	25,0
2	24,6	12	25,3	22	24,3
3	24,9	13	25,7	23	24,7
4	25,0	14	25,1	24	24,9
5	25,3	15	24,9	25	25,0
6	25,7	16	25,0	26	25,1
7	24,3	17	25,1	27	25,2
8	24,4	18	24,9	28	25,1
9	24,7	19	24,8	29	25,0
10	24,9	20	25,2	30	24,7

¿CÓMO SE UTILIZA?

2. Definir el recorrido R de esos datos:

$$R = X_{\max} - X_{\min}$$

donde:

X_{\max} = valor máximo de los datos

X_{\min} = valor mínimo de los datos.

EJEMPLO

En nuestro ejemplo:

$$R = X_{\max} - X_{\min} = 25,7 - 24,3 = 1,4$$

donde: $X_{\max} = 25,7$ mm y $X_{\min} = 24,3$ mm.

¿CÓMO SE UTILIZA?

3. Las *clases* son los intervalos en los que se dividen los valores de los datos. Establecer el número de clases **K**, que viene dado por la tabla siguiente:

Nº de datos (n)	Nº de clases (K)
20-50	6
51-100	7
101-200	8
201-500	9
501-1.000	10
Más de 1.000	11-20

Relación entre nº de datos y nº de clases.

EJEMPLO

En nuestro ejemplo, según la tabla anterior, para 30 muestras, $k=6$.

Nº de datos (n)	Nº de clases (K)
20-50	6

¿CÓMO SE UTILIZA?

4. Determinar la amplitud del intervalo de cada clase **h**, teniendo en cuenta que todas ellas tendrán la misma amplitud de intervalo y que no habrá solapamiento entre ellas.

Viene dado por la fórmula:

$$h = \frac{R}{K}$$

EJEMPLO

En nuestro ejemplo:

$$h = \frac{R}{K} = \frac{1,4}{6} = 0,23\bar{3} \cong 0,233$$

¿CÓMO SE UTILIZA?

5. Determinar el número de datos incluido en cada clase, es decir, la frecuencia de cada clase. Comprobar que el número total de datos es igual a la suma de la frecuencia de cada clase.

EJEMPLO

En nuestro ejemplo:

Clases	Valores límite de las clases	Chequeo	Frecuencias
1ª	24,3 - 24,533	///	3
2ª	24,534 - 24,767	////	4
3ª	24,768 - 25,001	#### ### /	11
4ª	25,002 - 25,235	### //	7
5ª	25,236 - 25,469	///	3
6ª	25,470- 25,7	//	2

¿CÓMO SE UTILIZA?

6. Trazar dos ejes, con las escalas siguientes:

Ordenadas: frecuencias.

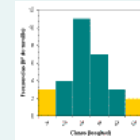
Abscisas: clases.

Una vez trazados los ejes, se llevan al gráfico del histograma los valores obtenidos para cada clase, trazando los correspondientes rectángulos que tendrán de base la amplitud de cada clase y de altura la frecuencia de las mismas.

Dibujar los límites de las especificaciones de la característica estudiada (si las hubiera).

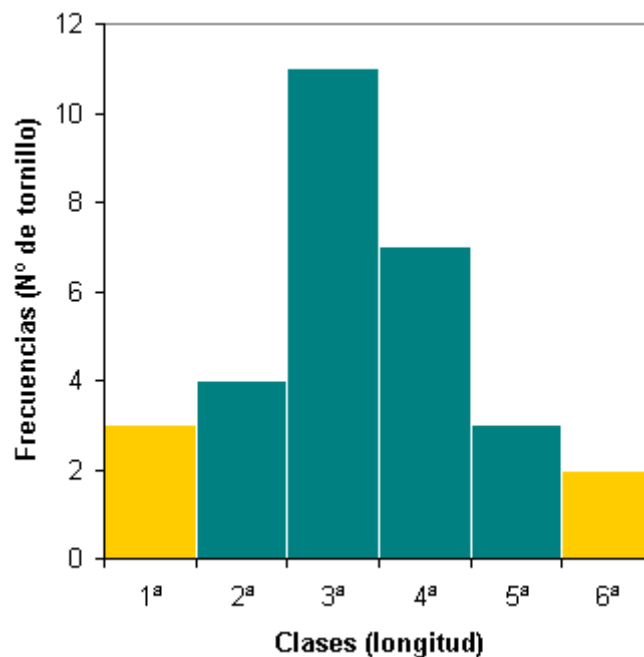
EJEMPLO

En nuestro ejemplo se representan gráficamente los datos: el eje X corresponde a las clases, es decir, a la longitud, y el eje Y corresponde a las frecuencias, es decir, al número de tornillos.



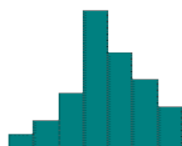
A continuación esta la imagen ampliada

Las muestras incluidas en las clases 1ª y 6ª corresponden a tornillos cuya longitud no cumple las especificaciones ya que están fuera del intervalo [24'6, 25'4].



¿CÓMO SE INTERPRETA?

Forma de campana o distribución normal: Simétrica, con el **pico** en el centro del rango de datos. Es la distribución más habitual de los datos de un proceso. No obstante, su aparición no significa que el funcionamiento del proceso sea correcto.



A continuación esta la imagen 1 ampliada

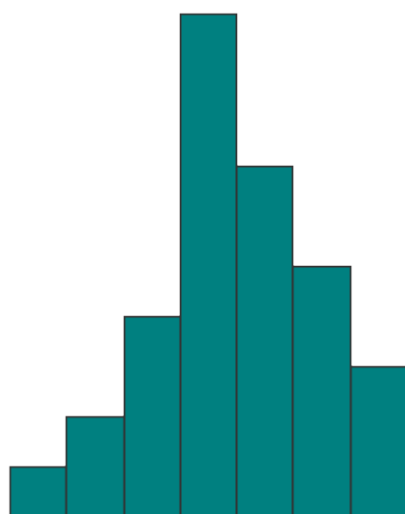
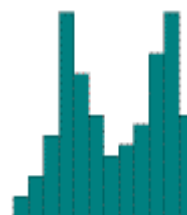


IMAGEN 1

Doble pico o Distribución bimodal: Tiene un **valle** en el centro del rango de los datos con dos **picos** en ambos lados. Esta es, en general, una Distribución de dos Distribuciones normales y sugiere la existencia de datos de dos procesos diferentes.



A continuación esta la imagen 2 ampliada

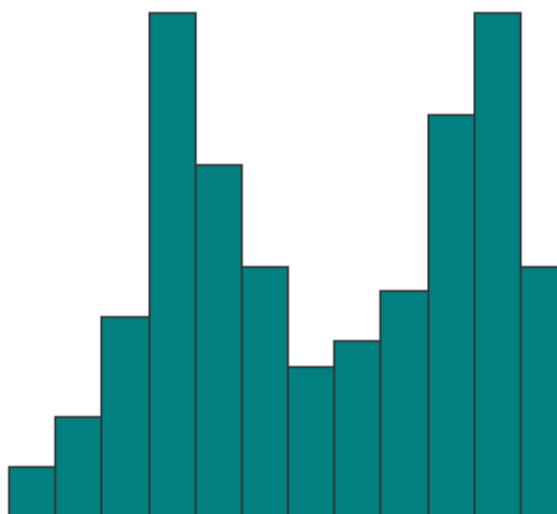


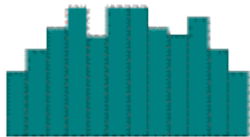
IMAGEN 2

Pico: En un histograma corresponde al valor modal o de mayor frecuencia.

Valle: En un histograma bimodal, es decir, con dos picos, el valle corresponde a la frecuencia menor entre los dos picos.

¿CÓMO SE INTERPRETA?

Meseta o distribución rectangular: Es una distribución uniforme con altos y bajos alternándose, donde no existe ningún [pico](#) claramente diferenciado. Sugiere la existencia de datos de varios procesos diferentes o se han producido errores de medición.



A continuación esta la imagen 1 ampliada

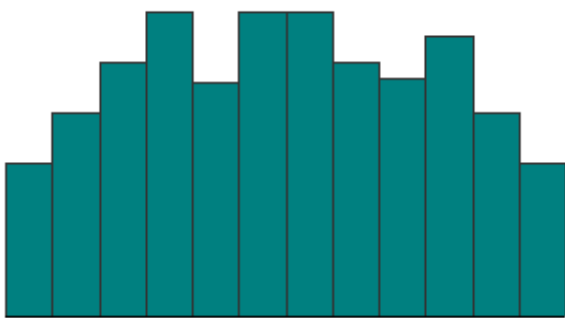
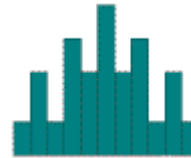
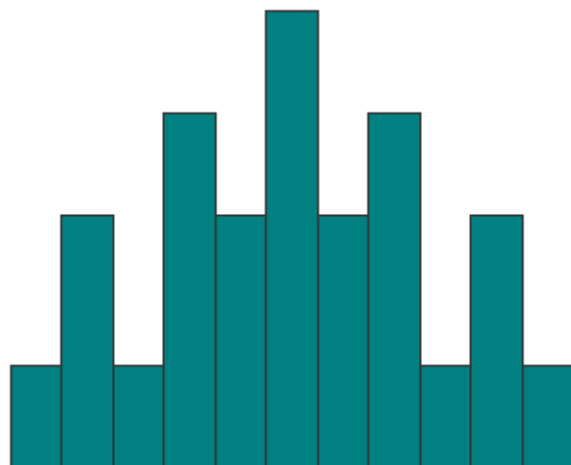


FIGURA 1
FIGURA 2

Peine: Alternan valores altos y bajos de forma regular. Refleja normalmente errores de medida.

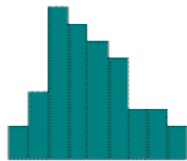


A continuación esta la imagen 2 ampliada



¿CÓMO SE INTERPRETA?

Sesgada: Es de forma asimétrica, con el [pico](#) descentrado respecto al rango de datos y la curva termina bruscamente en un lado y suavemente en el otro. Esta distribución puede ser debida a un proceso que se ha ajustado demasiado a uno de los extremos del [rango](#) permitido.



A continuación esta la imagen 1 ampliada

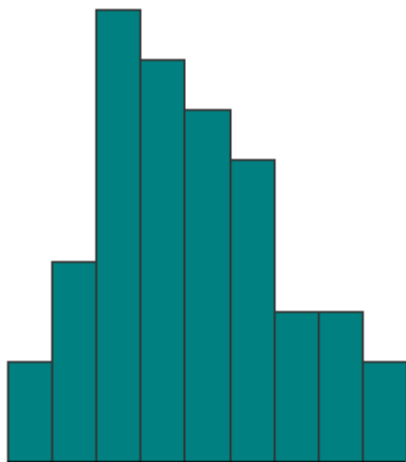
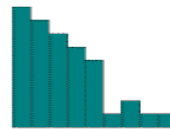


FIGURA 2

Truncada: Es una distribución cortada en uno o en dos de sus extremos. Normalmente se refiere a procesos con distribución normal, donde parte de la distribución se ha eliminado, por ejemplo, mediante una [inspección 100%](#). También puede aparecer debido a una mala elección del número de [clases](#).



A continuación esta la imagen 2 ampliada

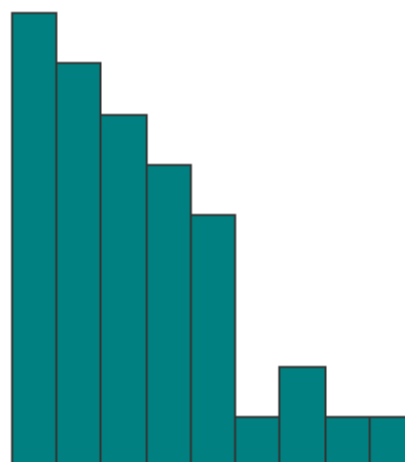


FIGURA 1

Rango: Véase recorrido.

Recorrido: Medida más sencilla de la dispersión. Diferencia entre el valor más alto y el más bajo de una característica estudiada.

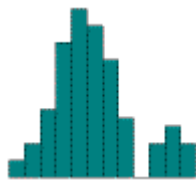
Inspección 100%: Inspección en la que se examinan todos y cada uno de los elementos de un conjunto.

Clase: Referida a características definidas por atributos: grupo de productos, establecido arbitrariamente, que tienen algunos atributos comunes. Referida a características definidas por variables: grupo establecido arbitrariamente que comprende una gama de valores de dicha variable.

Especificación: Documento que establece las características de un bien o servicio, tales como los niveles de calidad, el funcionamiento o comportamiento, la seguridad o las dimensiones.

¿CÓMO SE INTERPRETA?

Pico aislado: Con un grupo de datos pequeño separado de la distribución normal. Este pequeño grupo indica alguna anomalía del proceso, algo que no ocurre de forma regular.



A continuación esta la imagen 1 ampliada

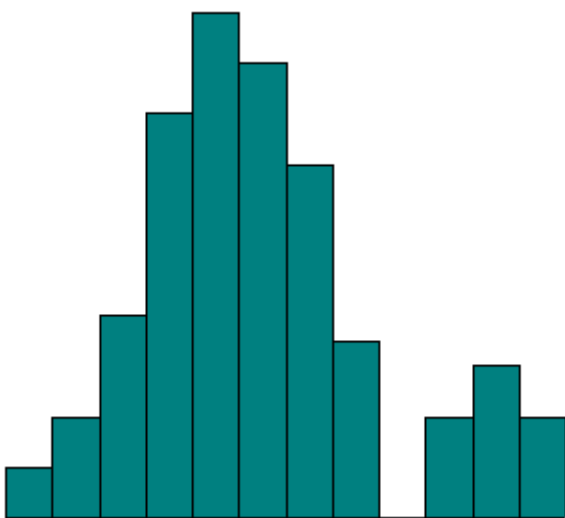
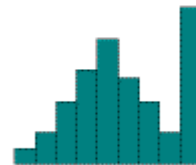


FIGURA 1

Pico extremo: Tiene un gran pico unido a una distribución normal. Indica normalmente un registro incorrecto de datos; por ejemplo, se han registrado valores fuera del rango aceptable como si estuvieran justo en el límite aceptable de la especificación.



A continuación esta la imagen 2 ampliada

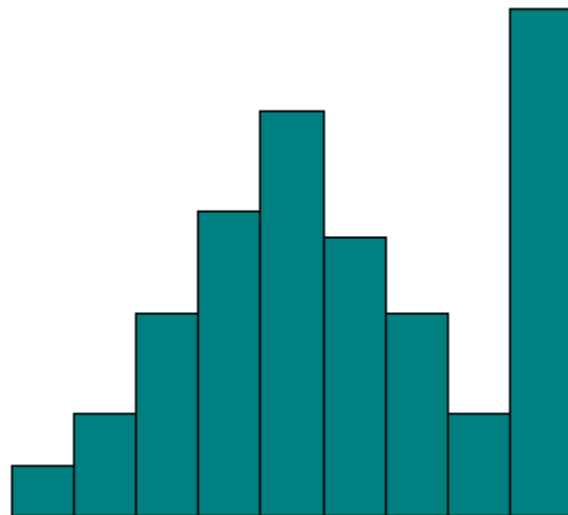


FIGURA 2

----- 00000 -----