

HISTOGRAMAS

1.- INTRODUCCIÓN

Este documento sirve de guía para la construcción de Histogramas y permite profundizar, más allá de su propia significación matemática, en el análisis del funcionamiento y la eficacia de procesos y procedimientos, a través de su información sobre la frecuencia de aparición de diferentes sucesos dentro de los mismos.

2.- OBJETIVO Y ALCANCE

Definir las reglas básicas a seguir para la construcción e interpretación de los Histogramas, resaltando las situaciones en que pueden, o deben, ser utilizados.

Es de aplicación a todos aquellos estudios en que es necesario analizar la pauta de comportamiento de un determinado fenómeno en función de su frecuencia de aparición.

Su utilización será beneficiosa para el desarrollo de los proyectos abordados por los Equipos y Grupos de mejora y por todos aquellos individuos u organismos que estén implicados en la mejora de la calidad.

Además, se recomienda su uso como herramienta de trabajo dentro de las actividades habituales de gestión.

3.- RESPONSABILIDADES

a) Grupo de trabajo o persona responsable del estudio:

- Recoger los datos.
- Seguir las reglas que se señalan en el procedimiento para la construcción del Histograma y para su correcta interpretación.

b) Dirección de Calidad:

- Asesorar, a quien así lo solicite, en las bases para la construcción y utilización de los Histogramas.

4.- DEFINICIONES / CONCEPTOS

4.1.- VARIABILIDAD

Definición

Campo de variación en los valores numéricos de una magnitud.

Concepto

Generalmente en los procesos de producción y de prestación de servicios es imposible mantener todos los factores que influyen en el resultado final, constantemente en el mismo estado.

Este hecho da lugar a que las características representativas del producto o servicio final presenten una determinada variación:

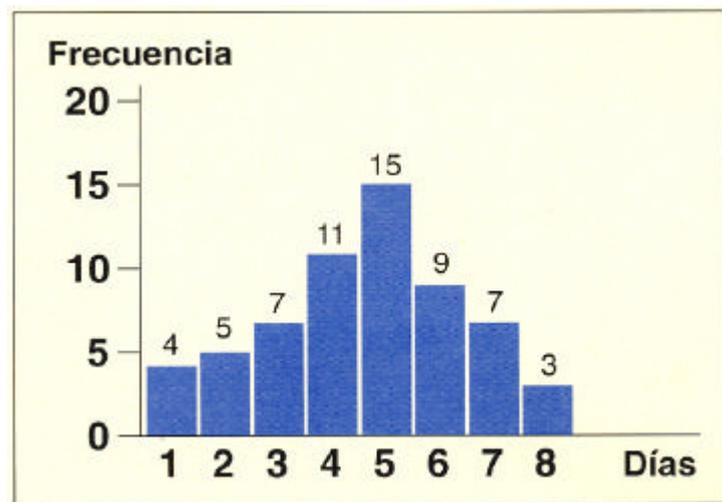
- *El tiempo de viaje para un determinado trayecto presenta diferencias de un día a otro debido a la variación de las condiciones de circulación, las condiciones climáticas, el número de viajeros, etc.*
- *Los ejes que produce una máquina tienen diferente diámetro dentro del mismo lote debido a pequeñas variaciones en las condiciones de la materia prima, a holguras de los elementos móviles, al desgaste de la herramienta, etc.*
- *El plato que prepara un cocinero tiene diferente gusto en diferentes ocasiones debido a variaciones en el peso de los condimentos utilizados, en el tiempo de cocción, etc.*

4.2.- HISTOGRAMAS

Definición

Es un resumen gráfico de los valores producidos por las variaciones de una determinada característica, representando la frecuencia con que se presentan distintas categorías dentro de dicho conjunto.

Ejemplo: Días pasados hasta contestar una reclamación



Características principales

A continuación se comentan una serie de características que ayudan a comprender la naturaleza de la herramienta.

Síntesis

Permite resumir grandes cantidades de datos.

Análisis

Permite el análisis de los datos evidenciando esquemas de comportamiento y pautas de variación que son difíciles de captar en una tabla numérica.

Capacidad de comunicación

Permite comunicar información de forma clara y sencilla sobre situaciones complejas.

4.3.- ESTRATIFICACIÓN

Definición

Separación de un conjunto de datos en diferentes grupos o categorías, de forma que los datos pertenecientes a cada grupo comparten unas características comunes que definen la categoría.

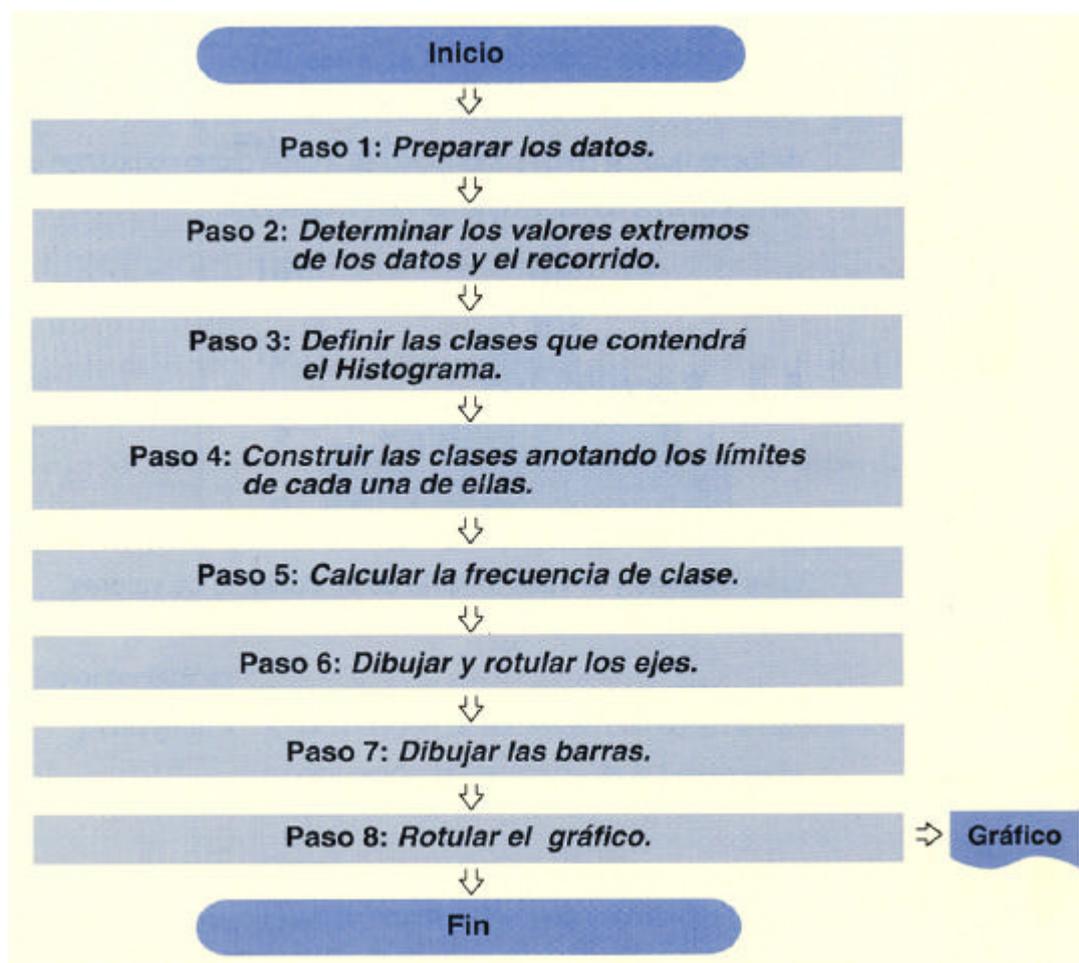
4.4.- RECORRIDO

Definición

Medida de la dispersión, correspondiente a la diferencia entre el valor máximo y el valor mínimo de un conjunto de valores.

5.- PROCESO

5.1.- DIAGRAMA DE FLUJO



5.2.- CONSTRUCCIÓN

Paso 1: Preparación de los datos

Como en todas las herramientas de análisis de datos, el primer paso consiste en recoger estos de forma correcta o asegurarse de la adecuación de los existentes.

Los datos deben ser:

- Objetivos: Basados en hechos, no en opiniones.
- Exactos: Debemos asegurarnos que la variabilidad en el proceso de recogida de datos (variabilidad de la medida) no desvirtúa la variabilidad del proceso en estudio.
- Completos: Se debe registrar toda la información relevante asociada a cada toma de datos (máquina, hora del día, empleado, etc) en previsión de los diferentes análisis que pueden ser necesarios.
- Representativos: Deben reflejar todos los diferentes hechos y circunstancias que se producen en la realidad.

Paso 2: Determinar los valores extremos de los datos y el recorrido

Identificar en la tabla de datos originales el valor máximo, el valor mínimo y el recorrido ($R = V_{\max} - V_{\min}$).

Ejemplo: Datos sobre la cantidad exacta de café contenido en paquetes de 250 gramos. (120 unidades medidas)

257	255	249	248	258	251	252	249	251	249
248	254	250	249	248	250	252	253	252	250
243	251	247	249	246	250	247	243	250	251
249	250	255	250	254	249	246	249	256	246
250	252	253	251	256	247	255	250	243	244
251	252	246	248	247	252	251	252	246	255
248	247	249	250	252	253	252	248	249	249
247	256	251	252	252	251	251	250	257	246
245	254	252	252	250	248	248	251	248	257
249	246	250	253	251	251	254	251	244	245
250	248	250	247	254	250	253	253	251	252
251	251	247	250	255	250	251	249	247	250

Recorrido total = máximo - mínimo = 258 grs. = 15 grs.

Paso 3: Definir las "clases" que contendrá el Histograma

Clases: Son los intervalos en que se divide la característica sobre la que se han tomado los datos. El número de clases es igual al de barras del Histograma.

a) Definir el número de clases que debe tener el Histograma según la tabla siguiente:

Numero recomendado de clases en un histograma

Número de datos	Número de clases recomendado
20* - 50	6
51 - 100	7
101 - 200	8
201 - 500	9
501 - 1000	10
Más de 1000	11 - 20

El mínimo para un histograma son 40 datos. Pueden darse menos si el histograma original ha sido estratificado.

Dependiendo del número de clases en que agrupemos los datos perderemos más o menos información tratando de identificar la pauta de comportamiento. La tabla anterior es un buen compromiso entre la máxima sencillez de análisis y la mínima pérdida de información.

b) Obtener la amplitud del intervalo de cada clase.

Todas las clases tendrán el mismo intervalo.

No habrá solapamiento entre distintas clases.

La amplitud aproximada del intervalo se halla dividiendo el recorrido por el número de clases.

Esta amplitud se redondea posteriormente a un número o cifra decimal conveniente para el manejo de las clases y la graduación del eje horizontal del Histograma (1, 2, 5, 10, etc).

En el ejemplo de los paquetes de café hay 120 datos, necesitamos entonces aproximadamente 8 clases para el histograma.

Amplitud aproximada de cada clase $15\text{gr.}/8=1.875\text{ grs.}$

Amplitud elegida como conveniente en este caso: 2 grs.

Paso 4: Construir las clases anotando los límites de cada una de ellas

Los límites de la primera clase incluirán el valor mínimo de los datos.

Para evitar que algunos datos coincidan con los límites de los intervalos, definir éstos de forma que tengan una cifra más detrás de la coma.

Si, por ejemplo, los datos tienen dos cifras detrás de la coma (3,55; 3,83; 3,64; 3,73; 3,78, etc), se definirán las clases hasta la tercera cifra detrás de la coma (3,545-3,555; 3,555-3,565, etc).

Si se obtiene una clase más o menos respecto del número recomendado, debido al redondeo posteriormente efectuado, no existe deterioro ni en la sencillez ni en la información.

Ejemplo: Como el valor menor en nuestros datos es 243 grs., empezaremos el primer intervalo en 242.5 grs. y construiremos ocho clases con 2 grs. de amplitud

Clase	Intervalo	Clase	Intervalo
1	De 242,5 a 244,5	5	De 250,5 a 252,5
2	De 244,5 a 246,5	6	De 252,5 a 254,5
3	De 246,5 a 248,5	7	De 254,5 a 256,5
4	De 248,5 a 250,5	8	De 256,5 a 258,5

Paso 5: Calcular la frecuencia de clase

Determinar el número de datos que están incluidos en cada una de las clases (frecuencia de clase).

El recuento se hará de la siguiente forma:

Empezar con el primer dato de la lista e identificar la clase en la cual está incluido. Señalar para dicha clase, un "palote". Repetir el mismo proceso para cada dato del conjunto.

Para facilitar el recuento final se dibujan los "palotes" en grupos de cinco, cuatro verticales y el quinto cruzándolos. La suma de los "palotes" marcados para cada clase corresponde a la frecuencia de la misma.

Comprobar que el número total de datos es igual a la suma de las frecuencias de cada clase.

Ejemplo

Límites de la clase	Recuento	Total
242,5 - 244,5		5
244,5 - 246,5		9
246,5 - 248,5		19
248,5 - 250,5		32
250,5 - 252,5		28
252,5 - 254,5		15
254,5 - 256,5		8
256,5 - 258,5		4
		120

Paso 6: Dibujar y rotular los ejes

El eje vertical representa las frecuencias, por tanto en él se rotularán números naturales, dependiendo su valor y escala del número de datos que se han tomado.

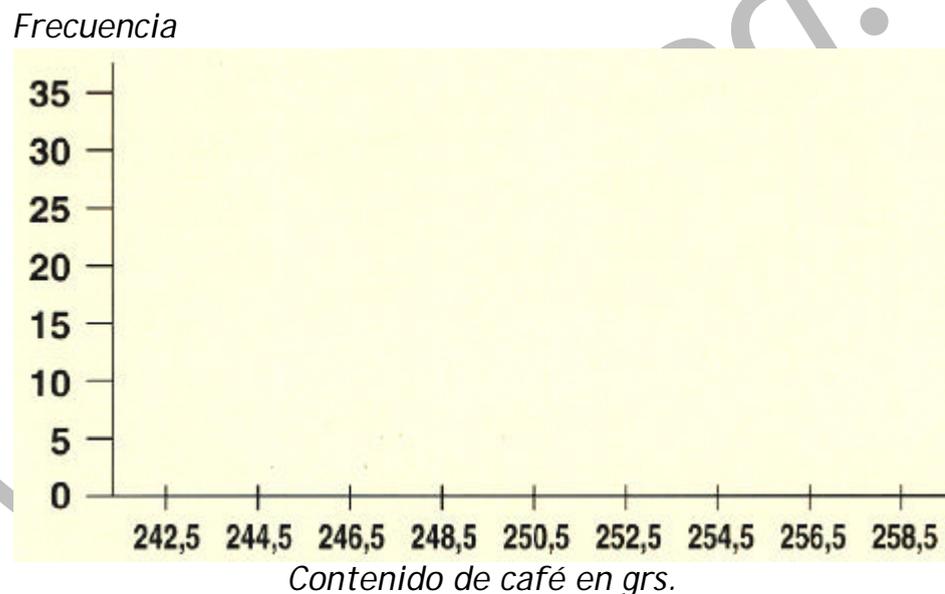
El eje horizontal representa la magnitud de la característica medida por los datos.

Este eje se divide en tantos segmentos iguales como clases se hayan definido.

Rotular los límites de los intervalos de clase.

Rotular el eje con la característica representada y las unidades de medida empleadas.

Ejemplo



Paso 7: Dibujar el Histograma

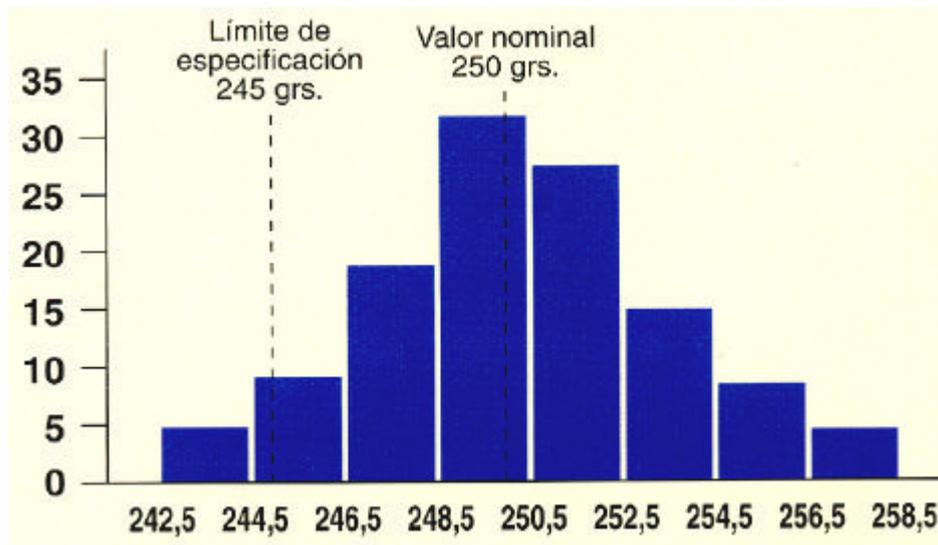
Dibujar las barras verticales correspondientes a cada clase. Su base está situada en el eje horizontal y su altura corresponderá a la frecuencia de la clase representada.

Paso 8: Rotular el Gráfico

Cuando proceda, poner el título, las condiciones en que se han recogido los datos, los límites de tolerancia nominales, etc. Estas notas ayudan a los demás a interpretar el gráfico y sirven de recordatorio de la fuente de los datos.

Ejemplo: Resultados de las mediciones del contenido de café (120 paquetes medidos)

Frecuencia



Contenido de café en grs.

5.3.- INTERPRETACIÓN

5.3.1.- Identificación e interpretación de las pautas de variación

Uno de los propósitos del análisis o interpretación de un Histograma es identificar y clasificar la pauta de variación del conjunto de datos estudiado (valor medio, recorrido, forma) y elaborar una explicación admisible y relevante para dicha pauta, que relacione la variación con el proceso o fenómeno en estudio.

El resultado de este análisis es una teoría sobre el funcionamiento del proceso o sobre la causa del problema que se está investigando.

Por ser una teoría es necesario confirmarla o rechazarla, recogiendo otros datos que nos den información más específica sobre dicha teoría.

La experiencia y habilidad del grupo de trabajo en la interpretación son fundamentales en la utilización de esta herramienta, puesto que no existen reglas fijas que se puedan utilizar para explicar de forma precisa las pautas de variación en cualquier situación.

Los equipos de trabajo deben profundizar en el conocimiento del proceso en estudio para utilizar esta herramienta de forma eficaz.

A continuación se presentan pautas de variación típicas que pueden ayudar a clasificar Histogramas y consejos generales sobre posibles explicaciones a las mismas.

Pautas típicas de variación:

Distribución en forma de campana

Forma simétrica con un pico en la mitad del recorrido de los datos.

Es la distribución natural, habitual para, los datos de gran cantidad de procesos. Por esta circunstancia se llama Distribución Normal.

La desviación respecto a esta forma puede indicar la existencia de problemas o influencias externas al proceso.

Sin embargo la forma de campana no asegura, por sí misma y sin analizar su valor medio y el recorrido de los datos, que el proceso funcione de forma satisfactoria.

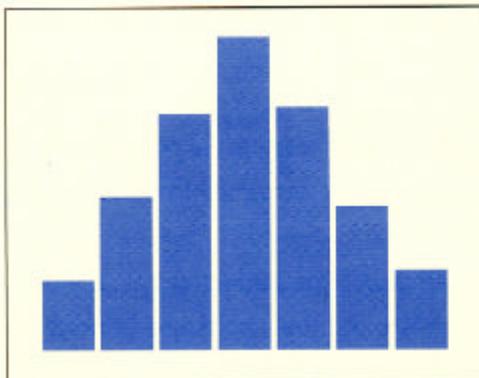
Distribución con doble campana o con doble pico

Un marcado valle en el centro de la distribución con picos a ambos lados.

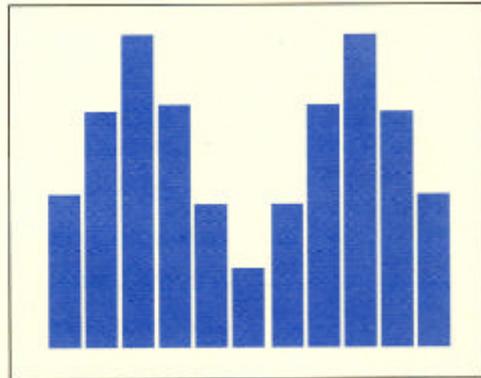
Esta forma, generalmente, es la combinación de dos distribuciones y sugiere la presencia de dos procesos distintos.

Deberán ensayarse varios esquemas de estratificación para separar los distintos procesos.

Distribución en campana



Distribución con doble pico



Distribución plana

Una gran parte plana, sin ningún pico y con dos ligeras colas a los lados.

Esta forma puede ser el resultado de varias distribuciones en campana con sus centros distribuidos uniformemente a lo largo del recorrido de los datos.

Se deberán identificar los diferentes procesos que intervienen dentro del proceso básico.

Esta distribución es un caso típico de departamentos u organizaciones que no tienen el trabajo bien definido y cada cual lo hace "a su manera".

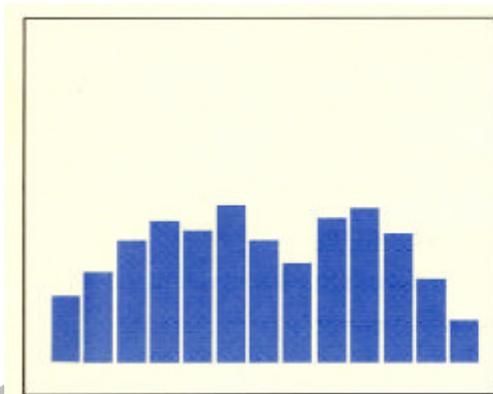
Distribución en peine

Valores altos y bajos se alternan de forma regular.

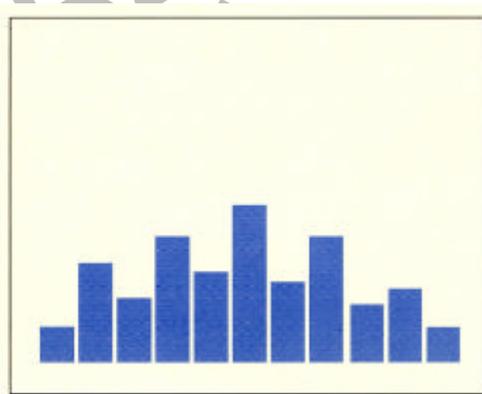
Esta pauta de variación es típica de errores de medición, errores en la forma de agrupar los datos para la construcción del Histograma o sesgos sistemáticos de redondeo.

En este caso revisar inicialmente los procesos de recogida de datos y construcción del Histograma.

Distribución plana



Distribución en peine



Distribución con un pico aislado

Como en el caso de la distribución de dos picos, esta forma sugiere la existencia de dos procesos distintos.

El proceso con el pico pequeño será una anomalía o deficiencia que no sucede a menudo o regularmente.

Se deben analizar las condiciones en que se presenta el pico menor tratando de estratificar los datos.

Estos picos unidos a distribuciones sesgadas o truncadas indican falta de eficacia en la eliminación de elementos defectuosos.

Distribución con un pico en el extremo

Un pico situado en un extremo de una distribución regular.

Esta forma se presenta cuando la cola de una distribución regular se ha cortado y acumulado en una sola categoría en el extremo del recorrido de los datos.

Suele indicar un registro poco cuidadoso o sesgado de los datos.

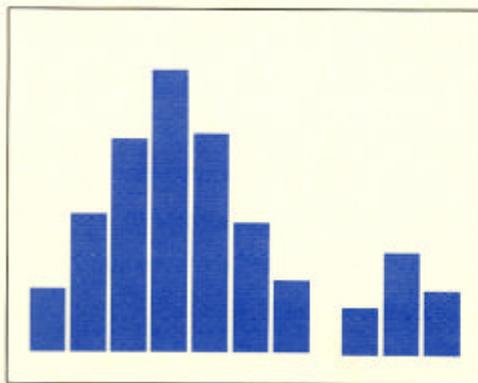
Ejemplo:

Se ha encargado al empleado responsable de contestar a las reclamaciones de los clientes que registre, durante un mes, el tiempo que se emplea en dar respuesta a las mismas.

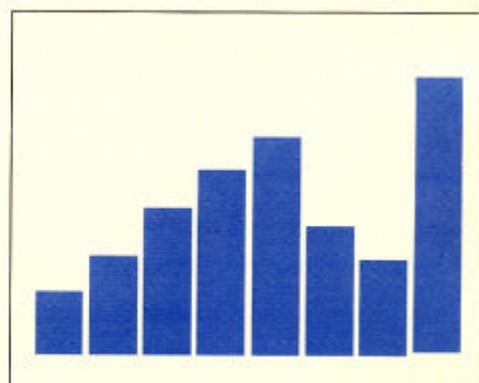
Para este proceso el procedimiento establecía un tiempo máximo de cinco días.

Para evitar reproches, el empleado registraba cinco días en todas aquellas ocasiones en que se superaba este límite, por lo que el Histograma correspondiente al tiempo de respuesta mostraba una "distribución con un pico en el extremo".

Distribución con pico aislado



Distribución con pico en extremo



Distribución sesgada o truncada

Su forma es asimétrica, con un pico descentrado dentro del recorrido de los datos, las colas descienden: bruscamente en un lado y suavemente en el otro. Esta distribución es típica de procesos con límites prácticos a un lado del valor nominal o a datos parciales de un proceso (distribuciones con parte de los datos suprimidos).

Ejemplo 1:

Al analizar el tiempo en que se tardan en cursar las órdenes de compra una vez recibidas, se observa que un gran porcentaje lo es en un día y el resto tardan dos, tres y hasta cuatro días debido a diferentes circunstancias.

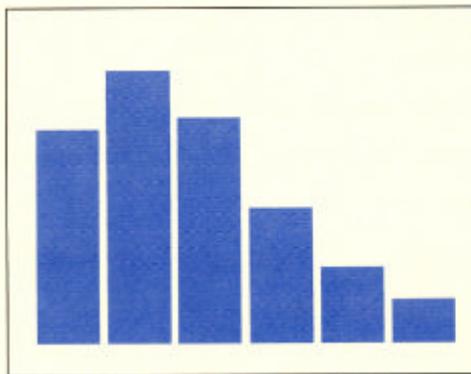
Al representar el Histograma de los tiempos requeridos para cursar dichas órdenes tendremos una "distribución truncada".

Ejemplo 2:

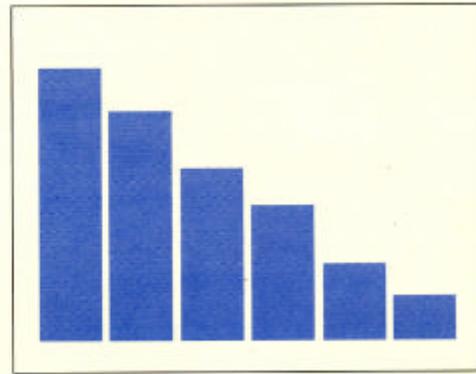
Al analizar el diámetro de los ejes de un pedido y construir el Histograma correspondiente, obtenemos una "distribución sesgada" cuyos extremos son similares a las tolerancias específicas para dicho diámetro.

Al contactar con los proveedores, estos nos comentan que han realizado una revisión completa del pedido (inspección 100%) para eliminar los ejes con diámetro fuera de tolerancias.

Distribución sesgada



Distribución truncada



5.3.2.- Comparación con los límites de aceptación o tolerancias

Es importante representar los límites de especificaciones u otros estándares, escritos o no, respecto de los resultados aceptables y considerarlos en la fase de análisis e interpretación de los Histogramas puesto que:

a) Los límites de especificaciones a menudo influyen en el proceso y sobre la pauta de variación de los datos.

El conocimiento detallado de estándares es esencial para la correcta interpretación de las distribuciones sesgadas, truncadas y con picos extremos.

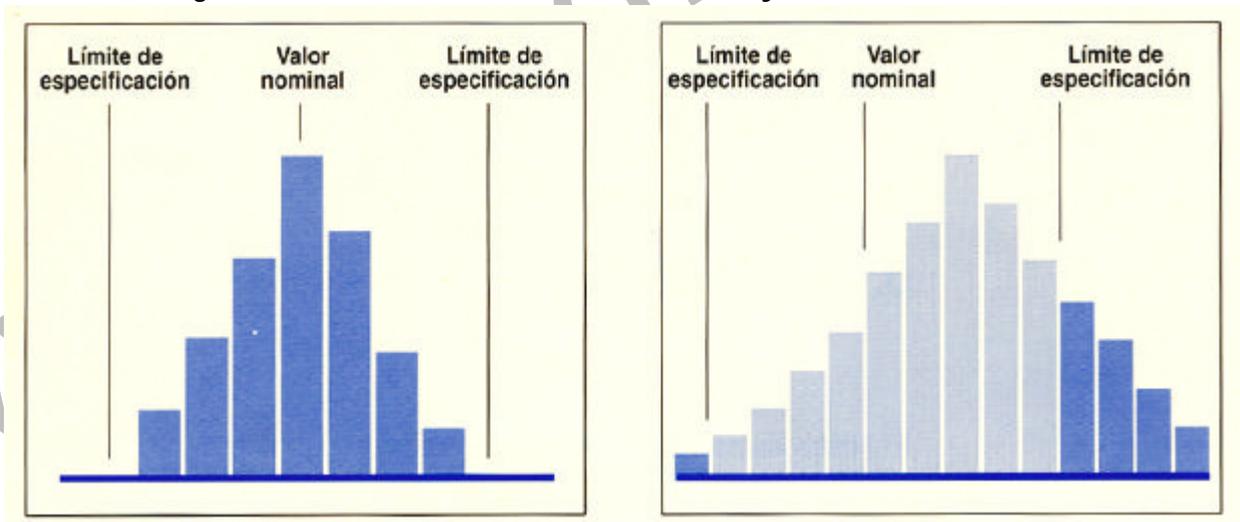
b) También se necesitan a menudo los límites para interpretar otras formas de Histogramas y decidir sobre los siguientes pasos a dar por el grupo de trabajo.

Incluso ante una distribución perfecta de campana, el grupo de trabajo debe plantearse cuestiones tales como:

¿Representa la distribución una actuación aceptable?

Aceptable, aunque las cosas cambien ligeramente. Hay margen de tolerancia

Inaceptable, es necesario centrar el proceso hacia el valor nominal, y reducir la variabilidad



5.3.3.- Posibles problemas y deficiencias de interpretación

a) Si los datos utilizados no son adecuados (sesgados, inexactos, anticuados, poco significativos, etc) las conclusiones no reflejarán la situación real.

b) Otra deficiencia en las conclusiones obtenidas del Histograma puede provenir de una muestra pequeña y poco representativa.

Como regla práctica se deberán recoger al menos cuarenta observaciones para cada uno de los Histogramas que se desee realizar.

En los casos en que esto no sea posible se deberán utilizar otros tipos de herramientas o consultar con un experto en estadística que diseñe un plan de muestreo y comprobación de hipótesis adecuado.

c) Aceptar las conclusiones del análisis como hechos.

La interpretación de un Histograma es una simple teoría y por tanto deberá ser confirmada posteriormente mediante el análisis adicional y la observación de los hechos reales.

5.4.- UTILIZACIÓN

El Histograma es una herramienta muy útil cuando un equipo se enfrenta con la tarea de analizar datos que presentan variaciones.

Utilización en las fases de un proceso de solución de problemas

En un proceso de solución de problemas hay dos puntos en los que la construcción y el análisis de Histogramas pueden ser muy útiles:

- Para la identificación de las Causas Raíz.

Se empieza generalmente el análisis con un Histograma de todos los datos del problema. El análisis de la pauta de variación de estos datos y generalmente posteriores estratificaciones conducen, paso a paso, a la identificación de las Causas Raíz.

- Para el seguimiento de los avances en las acciones de mejora de la calidad.

Se construyen con este fin Diagramas Antes-y-Después, teniendo cuidado de mantener la consistencia de la escala horizontal (= igual tamaño para intervalos iguales en los dos gráficos) y representando conjuntamente los dos Histogramas.

6.- ANEXOS

Ejemplo 1 Largas colas en las taquillas de una estación

Situación

En una estación se constituyó un equipo para examinar las quejas de los clientes sobre la lentitud en la venta y las largas colas en las horas punta.

Histograma y su análisis

El equipo midió los tiempos de las transacciones de cuatro empleados durante una hora punta típica y construyó con los datos obtenidos el siguiente Histograma:

Resultados de la recogida de datos:

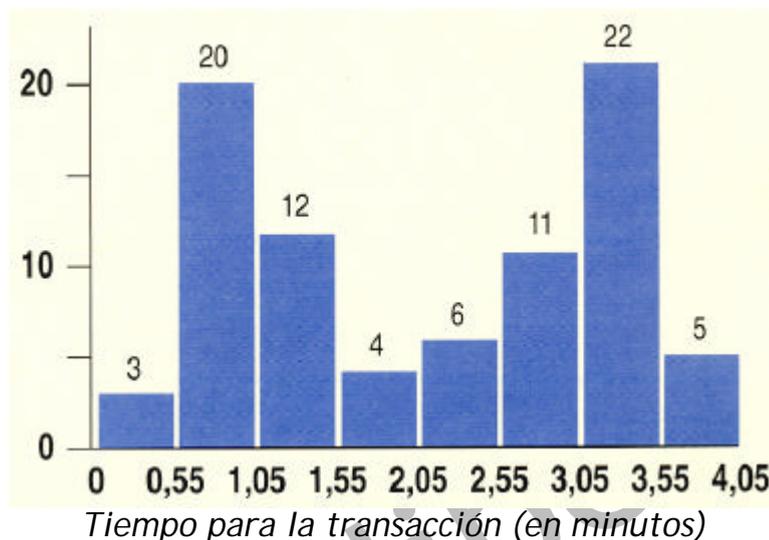
*Tiempo de realizar una transacción en horas punta
(83 transacciones - 4 empleados)*

0,8	2,5	3,5	3,8	3,3
0,7	2,3	3,4	3,9	3,2
0,9	2,2	4,0	4,0	0,8
0,7	1,7	3,8	3,5	
0,6	1,4	3,2	3,4	
1,2	1,2	3,3	3,1	
1,4	1,1	3,4	3,5	
3,2	1,0	2,8	3,3	
3,2	0,7	0,8	3,2	
3,3	0,8	0,9	2,8	
3,1	0,9	1,0	2,9	
3,5	0,6	0,6	1,7	
2,6	1,0	1,4	1,3	
2,8	0,5	1,2	1,2	
0,7	2,7	1,8	0,4	
1,0	2,8	1,9	0,7	
0,3	2,9	2,4	0,8	
1,4	3,3	2,5	2,4	
1,3	3,2	2,9	2,8	
1,4	3,2	3,0	3,3	

Histograma

*Tiempo para realizar una transacción en horas punta
(83 transacciones - 4 empleados)*

Frecuencia



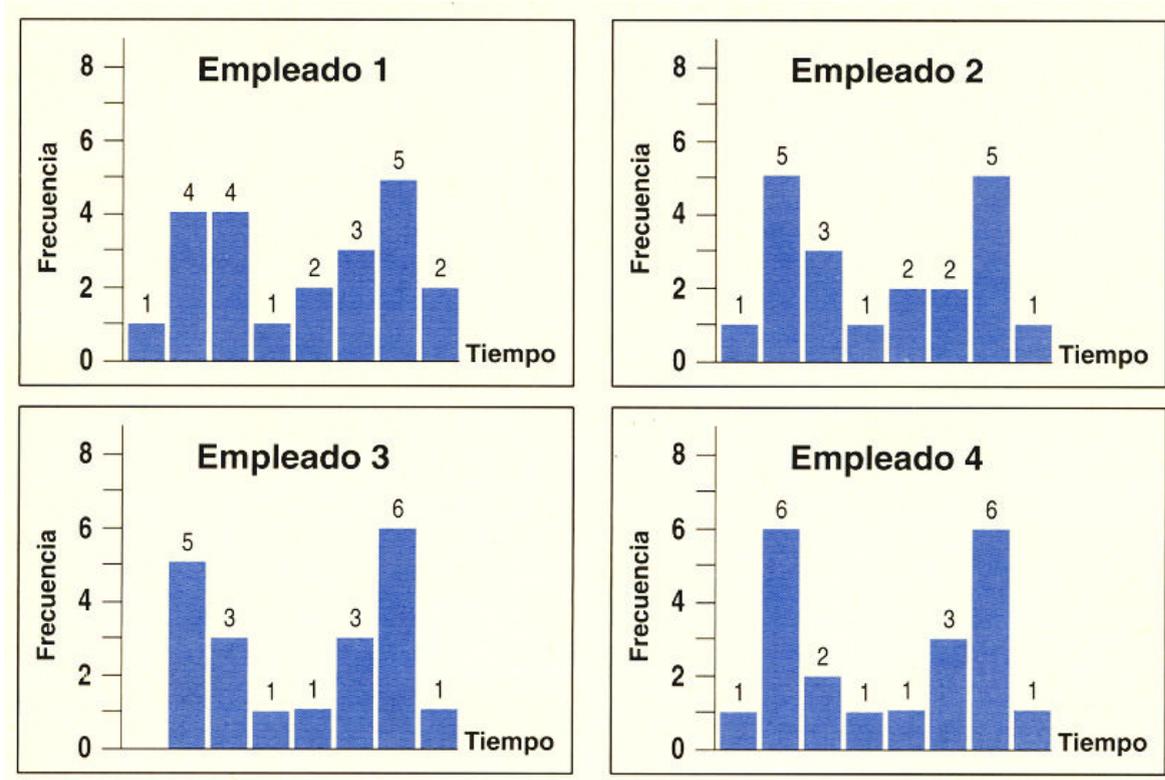
El Histograma muestra claramente una distribución con dos picos. En consecuencia el equipo formuló la teoría de que había dos procesos en marcha.

La primera teoría expuesta fue que los dos empleados nuevos eran simplemente más lentos que los más experimentados pero los Histogramas estratificados por empleados mostraron la misma distribución con dos picos, fuera cual fuera la experiencia.

En otras palabras, la estratificación por nivel de experiencia no sirvió para separar los datos en dos distribuciones en campañas distintas.

Estratificación por empleado

*Tiempo para realizar una transacción en horas punta
(83 transacciones)*



A continuación, el equipo decidió examinar más profundamente los tipos de transacciones que tenían lugar.

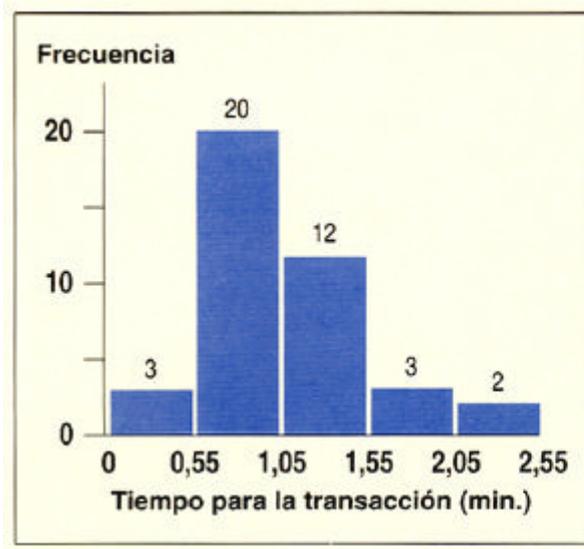
Esbozando diferentes Histogramas, finalmente fueron capaces de separar la distribución con dos picos, los tiempos de transacción cortos se asociaban a billetes para trayectos que no exigían combinaciones (operaciones de tipo A), mientras que los tiempos más largos estaban asociados a operaciones más complejas de búsqueda de combinaciones, información del cliente y emisión del billete (operaciones de tipo B).

El equipo había utilizado unos cuantos Histogramas para analizar los datos de forma más eficaz y para centrarse en las verdaderas causas de los retrasos en el servicio a los clientes.

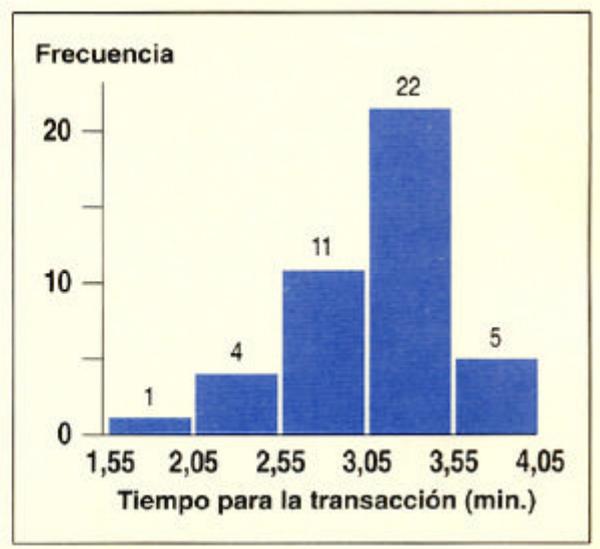
Estratificación por tipo de operación

Tiempo para realizar una transacción en horas punta
(83 transacciones - 4 empleados)

Operaciones de tipo A



Operaciones de tipo B



www.fundibeq.org

Ejemplo 2 Desarrollo de sistemas informáticos

Situación

Una empresa de sistemas informáticos estaba siendo sobrepasada por sus competidores que eran capaces de desarrollar un sistema y entregarlo al cliente en un promedio de 60 días, mientras que el promedio de la empresa era de 90 días.

El desarrollo de un sistema se realizaba en varios pasos y para cada uno de ellos existían estándares para los tiempos nominales y máximos.

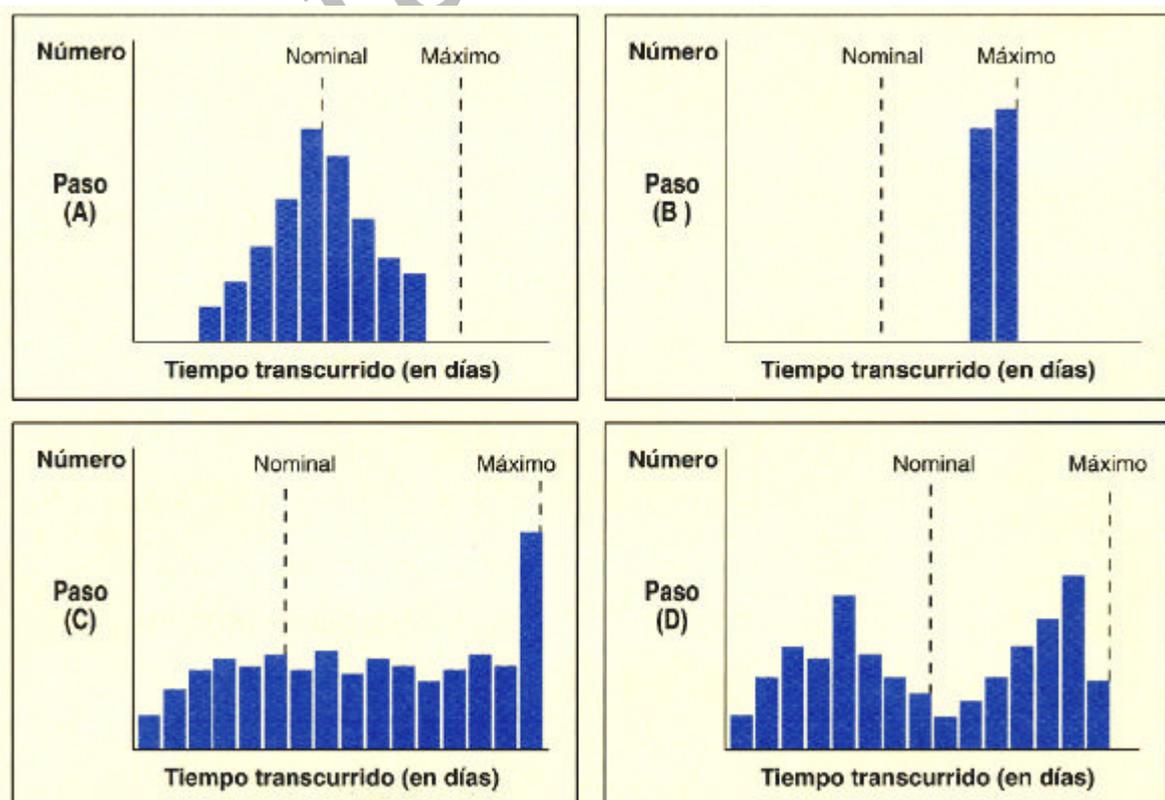
Pero el sistema no funcionaba: aunque no se había dado ningún caso en que un equipo rebasara los estándares globales para su paso, el tiempo global medio era todavía 30 días mayor que la meta de 60 días que los estándares se proponían.

Histogramas y su Análisis

Un equipo de mejora de la calidad recopiló los datos referentes a todos los sistemas desarrollados para clientes el año anterior y empezó construyendo Histogramas de los tiempos empleados en cada paso.

La figura siguiente muestra la distribución de los tiempos para cuatro de los pasos del proceso (A, B, C y D).

Distribución de los tiempos



Los histogramas incluyen todos los sistemas desarrollados en el año

La interpretación y conclusiones del equipo se resumen a continuación:

PASO (A) Un grupo bien conducido en que el trabajo progresó sin brusquedades hacia su terminación.

PASO (B) Un directivo o grupo moroso.

Un análisis ulterior reveló que aunque el trabajo se terminaba alrededor del tiempo fijado, (tiempo nominal), el equipo no lo entregaba hasta el último momento posible, por temor a tomar una decisión equivocada.

PASO (C) La forma plana de esta distribución indica que el grupo tiene en curso una gran variedad de procesos.

Un análisis ulterior reveló que la cantidad de trabajo que este grupo debía invertir variaba mucho de una persona a otra y de cliente a cliente. No había procesos estándar en el grupo ni una formación organizada para el nuevo personal.

Cada uno trabajaba en el desarrollo del sistema a su manera.

Se formó un equipo de mejora de la calidad para el examen de los procesos de trabajo de este paso, con el fin de reducir su variabilidad.

El segundo pico, justo dentro del límite máximo, tiene una explicación interesante: la distribución en realidad rebasaba en un cierto número de días el límite máximo. Pero cuando esto ocurría el supervisor del grupo registraba una fecha de terminación justo en el límite máximo, para no tener que dar explicaciones por el retraso.

PASO (D) En este paso existían dos procesos distintos, según el tipo de sistema. El grupo centró las acciones de mejora en aquellas circunstancias que llevaban a tiempos globales mayores.

7.- UTILIZACIÓN DE LA HERRAMIENTA

<i>La herramienta es muy útil para:</i>	<i>La herramienta es útil para:</i>
<ul style="list-style-type: none">- Determinación de causas- Evaluación de la solución implantada	<ul style="list-style-type: none">- Análisis de síntomas- Formulación de teorías sobre las causas- Evaluación de posibles soluciones- Diseño de soluciones y controles

8.- RELACIÓN CON OTRAS HERRAMIENTAS

<i>La herramienta está fuertemente relacionada con:</i>	<i>La herramienta está débilmente relacionada con:</i>
<ul style="list-style-type: none">- Hojas de Comprobación y Recogida de Datos- Estudios de Capacidad Potencial de Calidad	<ul style="list-style-type: none">- Cuadro de Mando Integral