

Histogramas

Un histograma es una gráfica con barras de la misma anchura, dibujadas una junto a la otra (sin espacios entre sí). La escala horizontal representa las clases de valores de datos cuantitativos, en tanto que la escala vertical representa frecuencias. Las alturas de las barras corresponden a los valores de frecuencia. Un histograma es básicamente la versión gráfica de una distribución de frecuencias.

Las barras en la escala horizontal se denotan con uno de los siguientes rótulos:

1. Fronteras de clase;
2. Marcas de clase; o
3. Límites inferiores de clase.

Las primeras dos opciones son técnicamente correctas, mientras que la tercera opción introduce un pequeño error. Se debe etiquetar ambos ejes con claridad.

El histograma es una gráfica de barras que permite describir el comportamiento de un conjunto de datos en cuanto a su tendencia central, forma y dispersión. El histograma permite que de un vistazo se pueda tener una idea objetiva sobre la calidad de un producto, el desempeño de un proceso o el impacto de una acción de mejora.

Construcción de un histograma.

Paso 1. Determinar el rango de los datos. El rango es igual a la diferencia entre el dato máximo y el mínimo.

Paso 2. Obtener el número de clases (NC). Existen varios criterios para determinar el número de clases (o barras). Sin embargo, ninguno de ellos es exacto. Se recomienda obtener de cinco a quince clases, dependiendo de cómo estén los datos y cuántos sean. Un criterio usado frecuentemente es que el número de clases debe ser aproximadamente igual a la raíz cuadrada del número de datos.

Paso 3. Establecer la longitud de clase (LC). La longitud de clase se establece de tal manera que el rango pueda ser cubierto en su totalidad por NC intervalos de igual magnitud. Así una forma directa de obtener la LC es dividiendo el rango entre el número de clases, $LC = R/NC$. Sin embargo, en ocasiones resulta más conveniente ampliar un poco el rango para que la longitud de clase esté expresada de manera más sencilla o en el mismo tipo de cifras que están los datos.

Histogramas

Paso 4. Construir los intervalos de clase. Los intervalos de clase resultan de dividir el rango (original o ampliado) en NC intervalos de longitud LC cada uno. El punto inicial para la primera clase puede ser el dato mínimo si no se amplió el rango. Si se amplió el rango, el punto inicial es un poco antes del mínimo.

Paso 5. Obtener la frecuencia de cada clase. Para obtener la frecuencia se cuentan los datos que caen en cada intervalo de clase. Cuando un dato coincide con el final de una clase y principio de la siguiente, entonces tal dato se incluye en esta última.

Paso 6. Graficar el histograma. Se hace una gráfica de barras en la que las bases de las barras sean los intervalos de clase y la altura sean las frecuencias de las clases.

Interpretación del histograma.

Cuando un histograma se construye de manera correcta y es resultado de un número suficiente de datos, en general más de 40, y estos son representativos de la población, proceso o problema, entonces lo que se aprecia en el histograma como tendencia central, variabilidad y comportamientos especiales será una información valiosa. Formalmente, observando un histograma se pueden contestar varias preguntas. Por ejemplo:

1. ¿Cuáles son las mediciones más comunes? Para ello hay que observar la barra o el grupo de barras más altos.
2. ¿Hay un comportamiento simétrico? ¿Hay sesgo? ¿Hacia qué lado? Para responder a estas preguntas basta observar el histograma.
3. ¿Cómo es la dispersión? Para contestar hay que observar a partir del grupo de barras más alto qué tan rápido disminuye la frecuencia de las demás barras.
4. ¿Está centrado el proceso? Con un tamaño de muestra grande es muy fácil ver mediante un histograma si un proceso está centrado o no, ya que basta observar la posición del cuerpo del histograma respecto a la calidad óptima y a las especificaciones.

Histogramas

5. ¿Cuántos picos hay? Cuando hay varios picos o agrupaciones de barras en un histograma, puede deberse a que el material procede de distintas cadenas de producción, de diferentes proveedores, han intervenido varios operadores o se han utilizado distintos instrumentos de medición sin sincronizar.

6. ¿Hay acantilados? Entre las posibles causas que motivan la presencia de un acantilado están: un lote de artículos previamente inspeccionados al 100% donde se excluyó a los artículos que no cumplen con alguna medida mínima o que exceden una medida máxima, problemas con el equipo de medición, errores en la inspección (cuando el inspector está predispuesto a no rechazar un artículo y observa que este casi cumple con los requisitos, registra la medida mínima aceptable). En general un acantilado es anormal y debe buscarse la causa del mismo.

7. ¿Hay datos aislados o raros? Un pequeño grupo de mediciones muy extremas o raras es fácilmente detectable en un histograma ya que aparece claramente aislado del resto. Un dato raro refleja una situación especial que se debe investigar, puede ser un error de medición o de dedo, pero también puede reflejar un comportamiento especial del proceso.

8. Estratificar. Cuando se obtienen datos que proceden de distintas máquinas, proveedores u operadores, puede encontrarse información valiosa si se hace un histograma por cada fuente (estratificar), con lo que se podrá encontrar la máquina o el proveedor más problemático.

Es recomendable que siempre que se analiza un histograma se dé respuesta por escrito a las siete preguntas anteriores y se contemple la posibilidad de estratificar, con lo que se sacará mayor provecho a la información y se irá adquiriendo la costumbre de proceder de acuerdo con un método.

Referencia:

Recuperado a través de Triola, M., 2004, Probabilidad y Estadística, Pearson Educación.
Gutiérrez Pulido Humberto, 1997, Calidad Total y Productividad, McGraw-Hill.