Serie Control de Proceso:

Distribución Normal
Prediciendo normalmente el futuro

Federico Salvador Wadsworth
Presidente Ejecutivo
Quality Consulting SA

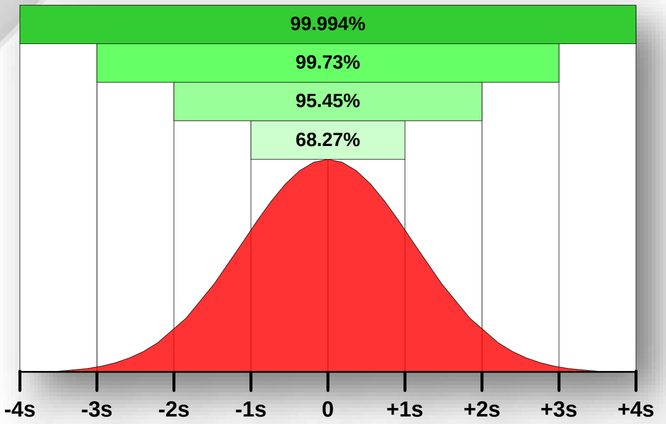
Un mundo ¿normal?

Vivimos en un mundo donde casi todo nos parece normal. La gente comenta que es “normal” que todo el mundo esté estresado, que la gente se divorcie, que la corrupción cunda, que la crisis no pare, que cada vez hayan más guerras, que los pobres en lugar de disminuir aumenten, que la capa de ozono siga desapareciendo…

Consideramos normal muchas cosas que antes se consideraban poco comunes, o que no existían. A veces cuando nos preguntan cómo estamos, decimos “normal”, por lo que el que escucha no sabe si debe condolerse de nosotros o felicitarnos, ya que la normalidad parece no ser siempre buena. Bueno… lo que pasa es que en realidad, normal no significa correcto, al menos no en la vida diaria. Pero ¿qué hay del mundo industrial? ¿es bueno tener un proceso normal?

La curva normal

Una de las herramientas más importantes de la calidad total para el estudio de los procesos productivos es la curva de distribución normal. De acuerdo a la experiencia y la estadística, en ese orden, puede comprobarse que la mayoría de procesos naturales y productivos se comportan de acuerdo a la distribución normal de probabilidades. Es posible asegurar que la estatura de los miembros de una población, el espesor del material que sale de una laminadora, el volumen envasado de un frasco de champú, responderá de alguna manera a la curva normal.



Esta curva, llamada también la campana de Gauss (ver el gráfico), permite analizar estadísticamente los procesos y predecir su comportamiento futuro.

Los parámetros de la distribución normal de probabilidades son la media denotada por μ:, que es estimada por la media aritmética, y la desviación estándar σ, estimada por la desviación estándar muestral s_{n-1}. En base a esto, podríamos señalar que una variable se comportará de acuerdo a la curva normal si posee las siguientes características:

1. La variable es continua, por lo que puede adquirir cualquier valor en la escala. Es decir debe ser medible y no contable. Por ejemplo: el número de unidades defectuosas de un proceso no responde a la curva normal ya que sólo puede adquirir los valores 0, 1, 2, 3 y así sucesivamente. El porcentaje defectuoso que se calcule en base a estos tampoco corresponderá a la curva normal. En cambio el espesor de una pieza, el volumen interior de una botella de gaseosa sí son variables continuas.
2. La distribución es simétrica con respecto a la media, es decir, existe igual probabilidad que los valores caigan por debajo de la media que por encima de ella. Algunas variables de proceso, como el tiempo de la espera en la cola de una ventanilla del banco, no dependen de la curva normal ya que la probabilidad que caigan por debajo de la media es casi el doble que la probabilidad que la excedan. Otra distribución de probabilidades será la adecuada.
3. La distribución es continua desde menos infinito a más infinito, sin embargo, la probabilidad que la variable alcance valores alejados de la media va decreciendo velozmente. El 68,27% de los valores se ubicarán en el rango de $\pm$ 1σ, el 95,45% dentro de $\pm$ 2σ y el 99,73% dentro de $\pm$ 3σ. Apenas 6 de cada 100,000 valores excederán la cuarta desviación estándar, 6 en 10 millones la quinta y 2 en 1,000 millones la sexta.
4. El área bajo la curva normal es igual a 1, como ocurre en todas las distribuciones de probabilidad sean estas continuas, como la normal, o discretas (cuyas variables sólo pueden adquirir valores enteros), como la binomial.
5. La media, la moda y la mediana de los datos coinciden en un mismo punto.

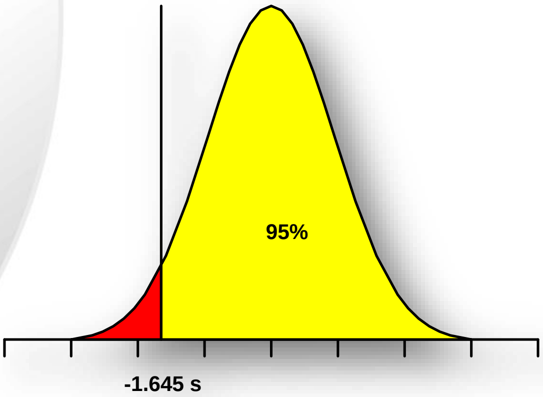
Método de predecir el futuro

Cuando una variable responde a la curva normal se denota de la siguiente manera: $x \sim N(\bar{\bar{X}}, s)$; lo que significa que la variable x se comporta de acuerdo a una curva normal de media X y de desviación estándar s. Para una curva normal dada puede integrarse la función de densidad entre dos valores determinados obteniéndose un área que representa la probabilidad que la variable se encuentre entre dichos límites. Este procedimiento requiere de una integración compleja que se simplifica con el uso de las tablas de distribución normal de z.

Toda curva normal puede ser transformada a una del tipo z ~ N (0,1), donde z (para un valor dado de x) es un estadístico de la forma:

$$z = \frac{(x - media)}{s}$$

Al obtener el valor de z equivalente a x es posible utilizar las tablas de distribución normal que permiten saber qué porcentaje o fracción de los valores se encuentran por debajo de dicho valor.



También las hojas de cálculo, MS Excel por ejemplo, con la función:

DISTR.NORM.ESTAND.N(z,VERDADERO)

permiten obtener el porcentaje bajo la curva para cualquier valor de z o directamente digitando la media y la desviación estándar. Vea un ejemplo gráfico (con la porción en rojo) para un valor de z=-1,645 que indicaría que un 5% están por debajo del flor de z y un 95% por encima de él.

De manera que: conociendo el promedio y desviación estándar de un proceso podemos determinar qué proporción de los valores se encontrarán dentro de los límites de especificación o por encima o por debajo de un valor cualquiera. Si un proceso se encuentra bajo control y se comporta en forma estable es posible predecir su comportamiento futuro a través de la distribución normal.

Veamos esto a través de un ejemplo.

Una aplicación práctica

Supongamos que el proceso de envasado de lubricante se realiza a través de boquillas controladas automáticamente. Una computadora registra el flujo de lubricante determinando el momento en que la boquilla se cierra. En teoría, el envase excede con alguna amplitud el nivel máximo esperado por lo que no se espera derrames. Sin embargo es común observar que los envases salen en alguna proporción manchados por el exceso de lubricante inyectado.

Se realizó un estudio para determinar el volumen promedio y la variación que provocaban las distintas boquillas. Luego de medir el volumen en 100 inyecciones consecutivas se obtuvo un promedio 1,023 galones con una desviación estándar de .016 galones. El envase tiene un volumen de 1.045 gal y se deseaba estimar de acuerdo a esto que proporción de los envases mostrarían rebalse en una producción futura. Asimismo la especificación de llenado es de .99 a 1.01 galones por envase, y es necesario determinar qué proporción de los envases estarán dentro de la tolerancia.

Calcularemos primero los valores de z correspondientes y que porcentaje de los valores se encuentran por debajo de cada valor:

Para X = 1,045 (el volumen máximo del envase):

$$z = \frac{(1{,}045 - 1{,}023)}{.016} = +1{,}375$$

Buscando el valor de z en tablas (o en Excel) se encuentra que corresponde a un 91,54%. Esto implica que un 8,46% de envases se rebalsarán en producción. Asimismo al calcular los valores de z para los límites de la especificación encontramos -2.06 y -.81. Esto corresponde a porcentajes de 1.96% y 20.83%.

Esto significa que 1.96% de los envases contendrán menos de .99 galones y 79.17% más de 1,01 galones; lo que deja un porcentaje de 18.87% en tolerancia.

Esta información puede resultar valiosa tanto para planificar la producción como para decidir las acciones correctivas que en este caso serían necesarias. Si se produjeran 30,000 galones al día, esto implicaría que 2,538 envases (30.000 por 8,46%) deberán ser limpiados. Si un hombre demora 10 segundos en limpiar cada envase se necesitaría 7,05 horas-hombre diarias para la limpieza de estos envases. Pero si se considera que, por el problema existente, es necesario revisar todos los envases y esto representara unos 5 segundos extras por envase, se requeriría otras 41,7 horas-hombre al día.

Suponiendo por otro lado que el galón de lubricante tenga un costo de US$ 4.05, el sobrecosto por volumen sería de algo más de 9 centavos por galón lo que reportaría una pérdida diaria de casi 2,800 US$, es decir, unos 850,000 US$ anuales. Interesante ¿verdad?

Un poco de cautela

La distribución normal no puede ser aplicada a todas las variables, como ya mencionamos, pero podría además ser erróneo emplearla en procesos que, a pesar de ser teóricamente normales, no estén bajo control. Si un proceso es afectado por causas que provocan variaciones asignables (que puedan ser asignadas a una causa), se convierte en un proceso impredecible. Por otro lado, el estudio de un proceso en estas condiciones podría llevarnos a conclusiones y predicciones erróneas.

Esto establece una íntima relación entre el concepto de cartas de control para variables y la curva normal, ya que sólo es posible predecir el comportamiento futuro de procesos que se encuentran en estado de control estadístico.

Curvas "normaloides"

La curva normal tiene una única función de densidad y por lo tanto una relación entre la amplitud y la altura de la curva. Posee también una simetría a la que hemos hecho referencia. Cualquier otra curva que no posea estas características es una especie de curva normal o "normaloide".

Es importante entonces que, cuando se estudia un proceso para determinar su grado de normalidad, y se evalúe aspectos como el sesgo y la kurtosis.

El sesgo (skewness) o asimetría se manifiesta como un "descentramiento" de la media con respecto a la amplitud o variación. Veamos la relación entre la media y la moda. Si la media es mayor que la moda (el valor más frecuente, aparecería como el pico en la curva) se considera que la curva tiene un sesgo a la derecha o positivo. Esto se manifestará como una curva con la "cola" derecha más larga. El caso contrario significará un sesgo negativo o a la izquierda. Se mide con un parámetro denominado Sk. Si es mayor que cero se considera a la derecha o positivo. Si Sk es igual a cero, la curva es simétrica con respecto a la media. Un valor aceptable de sesgo para poder predecir con la curva normal es ± .20.

La kurtosis en cambio es una medida de la relación existente entre la altura de curva y su amplitud. Una curva con Ku igual a cero indicará que se trata de una curva mesocúrtica o normal. Si el valor excede a cero significará que se trata de una curva leptocúrtica, es decir, aquella que tiende a concentrar más valores en la zona central. Lo contrario será una curva platicúrtica, o aquella con menor concentración de valores en la zona central.

Una kurtosis aceptable para predecir con la curva normal será de ± .20.

Conclusiones

La curva normal es una poderosa herramienta para estudiar y controlar los procesos manufactureros y administrativos. Posee una íntima relación con las cartas de control para variables y puede ser usada para procesos que, sin ser normales, pueden ser convertidos en tales en base a una transformación algebraica.

Tome en cuenta que no se puede controlar un proceso que no se conoce. El conocimiento del proceso se adquiere midiendo qué tan grande es la variación. Uno de los gurús de la calidad la define como: ausencia de variación. Conozca su proceso a través de la curva normal.

Yo también pienso que el proceso de una vida de calidad debe estar ausente de variación. Debemos vivir siempre en base a los principios correctos. Tome su tiempo para analizar qué tan variable es su vida frente a los principios divinos. Hay un pasaje bíblico que dice: "*Hay camino que al hombre le parece derecho; pero su fin es camino de muerte*".

No oscile entre lo que es correcto y lo que no lo es. Con la ayuda de Dios mantenga su vida en control.