

TEMA 6: Gráficos de Control por Variables

- 1 Introducción
- 2 Gráficos de control de la media y el rango
 - Función característica de operación
- 3 Gráficos de control de la media y la desviación típica
- 4 Gráficos de control de la media y la varianza
- 5 Gráficos de control modificados
- 6 Gráfico de control de sumas acumuladas
- 7 Gráficos de control basados en medias ponderadas
 - Gráfico de control de la media móvil
 - Gráfico de control de la media geométrica
- 8 Gráficos de control multivariantes

1. Introducción

Los gráficos de control por variables permiten estudiar la calidad de características numéricas. Proporcionan más información que los gráficos de control por atributos sobre el rendimiento del proceso y permiten procedimientos de control más eficaces. En particular, se obtiene más información sobre las causas que producen una situación fuera de control. Asimismo detectan mejor pequeñas variaciones del proceso. Los tamaños muestrales requeridos para un nivel de protección del proceso son menores.

Los gráficos de control por variables más usuales son los que controlan el valor medio y la variabilidad del proceso. Más concretamente, para el control de la variabilidad del proceso, estudiaremos los gráficos del rango, la desviación típica y la varianza. A partir de estos gráficos se obtiene una estimación de los parámetros del proceso, así como una aproximación de su capacidad o rendimiento.

Existen otros gráficos por variables que utilizan la información suministrada por varias muestras; se construyen a partir de una secuencia de estadísticos muestrales, en contraste con los gráficos anteriores, que tan sólo utilizan la información de una muestra. Estos gráficos permiten detectar pequeños cambios en la media del proceso, corrigiendo así la ‘incapacidad’ de los gráficos anteriores, para controlar pequeñas variaciones. En este tema se estudiarán dos ejemplos de estos gráficos: Los gráficos de sumas acumuladas y los gráficos de control basados en medias ponderadas. Por otra parte, para el estudio de

varias características relacionadas se introducirán los gráficos multivariantes basados en la distribución $\mathcal{T}^2$ de Hötelling.

2. Gráficos de control de la media y el rango

Supongamos que la característica de la calidad estudiada, X , se distribuye según una Normal con parámetros μ y σ . Es bien conocido que, para una muestra de tamaño n , la media muestral $\bar{X}$ se distribuye según una Normal con parámetros μ y $\frac{\sigma}{(n)^{1/2}}$.

En el caso en el que μ y σ sean parámetros conocidos, los límites del gráfico de control vienen dados por

$$\mu \pm 3 \frac{\sigma}{(n)^{1/2}}.$$

Cuando los parámetros poblaciones son desconocidos, se estiman a partir de m muestras piloto como sigue:

- Para la media μ se considera el estimador

$$\hat{\mu} = \bar{\bar{x}} = \frac{\bar{x}_1 + \dots + \bar{x}_m}{m}.$$

- Para la desviación típica σ se considera el estimador

$$\hat{\sigma} = \frac{R_1 \dots R_m}{m d_2} = \frac{\bar{R}}{m d_2},$$

donde R_j , $j = 1, \dots, m$, representan los rangos o amplitudes muestrales de las m muestras piloto y d_2 representa la media de la amplitud relativa A_R de una muestra normal (que se halla tabulada), siendo

$$A_R = \frac{R}{\sigma}.$$

Este estimador proporciona buenas aproximaciones de la desviación típica poblacional para tamaños muestrales pequeños.

Los límites de prueba para el gráfico de control de la media vienen dados entonces por

$$\bar{\bar{x}} \pm 3 \frac{\bar{R}}{d_2(n)^{1/2}} = \bar{\bar{x}} \pm 3 A \bar{R}.$$

La constante A se halla tabulada para diferentes tamaños muestrales.

Para el cálculo de los límites de control para el gráfico del rango se considerará de nuevo la amplitud relativa de una muestra normal, obteniéndose de la identidad

$$R = \sigma A_R,$$

y, por tanto, de la identidad

$$\sigma_R = \sigma \sigma_{A_R} = \sigma d_3,$$

la siguiente estimación:

$$\hat{\sigma}_R = \frac{\bar{R}}{d_2} d_3,$$

donde d_3 se halla también tabulada para diferentes tamaños muestrales. Los límites de prueba para el gráfico del rango se definen entonces como sigue:

$$\bar{R} \pm 3d_3 \frac{\bar{R}}{d_2},$$

o, equivalentemente,

$$\begin{aligned} LIC &= D_3 \bar{R} \\ LSC &= D_4 \bar{R}, \end{aligned}$$

siendo

$$\begin{aligned} D_3 &= 1 - 3 \frac{d_3}{d_2} \\ D_4 &= 1 + 3 \frac{d_3}{d_2}. \end{aligned}$$

Metodología. Si existen patrones no aleatorios en los gráficos de la media y el rango, se eliminan primero las causas atribuibles en el gráfico del rango, lo que producirá una mejora en el gráfico de $\bar{X}$. Bajo la hipótesis de normalidad, $\bar{X}$ y R son independientes. Por tanto, si esta hipótesis no se refleja en los gráficos de control de ambas variables, se debe revisar la hipótesis de normalidad.

2.1. Función característica de operación

Describe la idoneidad de los gráficos de $\bar{X}$ y R para detectar cambios en la calidad del proceso. Si se produce un cambio en la media del proceso de un valor μ_0 a un valor $\mu_1 = \mu_0 + k\sigma$, entonces,

$$\begin{aligned} \beta = P [LIC \leq \bar{X} \leq LSC / \mu_1 = \mu_0 + k\sigma] &= P \left[Z \leq \frac{LSC - (\mu_0 + k\sigma)}{\sigma/(n)^{1/2}} \right] \\ &- P \left[Z \leq \frac{LIC - (\mu_0 + k\sigma)}{\sigma/(n)^{1/2}} \right]. \end{aligned}$$

Puesto que $LIC = \mu_0 - 3 \frac{\sigma}{(n)^{1/2}}$ y $LSC = \mu_0 + 3 \frac{\sigma}{(n)^{1/2}}$, se obtiene

$$\beta = P [Z \leq 3 - k(n)^{1/2}] - P [Z \leq -3 - k(n)^{1/2}].$$

3. Gráficos de control de la media y la desviación típica

Se considera el estadístico S definido a partir de la raíz cuadrada de la cuasivarianza muestral,

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}.$$

Puesto que S no es un estimador insesgado de σ , siendo $ES = c_4\sigma$ con

$$c_4 = \left(\frac{2}{n-1} \right)^{1/2} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)},$$

los límites de control para el gráfico de σ , en términos del estadístico S vienen dados por

$$c_4\sigma \pm 3\sigma(1 - c_4^2)^{1/2},$$

o, equivalentemente,

$$\begin{aligned} LIC &= \sigma B_1 \\ LSC &= \sigma B_2, \end{aligned}$$

siendo

$$\begin{aligned} B_1 &= c_4 + 3(1 - c_4^2)^{1/2} \\ B_2 &= c_4 - 3(1 - c_4^2)^{1/2}, \end{aligned}$$

en el caso en que σ es conocida. Para σ desconocida se considera la extracción de m muestras piloto y se obtiene

$$\bar{S} = \frac{1}{m} \sum_{j=1}^m S_j,$$

denotando por S_j la raíz cuadrada de la cuasivarianza muestral de la muestra piloto j . Los límites de prueba vienen dados por

$$\bar{S} \pm 3 \frac{\bar{S}}{c_4} (1 - c_4^2)^{1/2}.$$

El gráfico para la media se define entonces en términos de los siguientes límites de control:

$$\bar{\bar{X}} \pm 3 \frac{\bar{S}}{c_4(n)^{1/2}}.$$

4. Gráficos de control de la media y la varianza

Puesto que los límites iniciales para el gráfico de la media serían

$$\bar{\bar{X}} \pm 3 \left(\frac{\sigma^2}{n} \right)^{1/2},$$

en el caso en que σ es desconocida se aproximan en términos de la media de las cuasivarianzas muestrales de las m muestras piloto,

$$\bar{S}^2 = \frac{\sum_{j=1}^m S_j^2}{m}.$$

Los límites de control de S^2 no pueden calcularse en términos de un múltiplo de la desviación típica de dicho estadístico, puesto que su distribución no es simétrica. Sin embargo, se pueden calcular los límites probabilísticos para S^2 teniendo en cuenta que

$$\begin{aligned} 1 - \alpha &= P \left[\chi_{1-\frac{\alpha}{2}, n-1}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\frac{\alpha}{2}, n-1}^2 \right] \\ &= P \left[\frac{\sigma^2 \chi_{1-\frac{\alpha}{2}, n-1}^2}{n-1} \leq S^2 \leq \frac{\sigma^2 \chi_{\frac{\alpha}{2}, n-1}^2}{n-1} \right], \end{aligned}$$

es decir, utilizando la estimación anterior $\bar{S}^2$ de la varianza se obtienen los límites de control

$$LIC = \frac{\bar{S}^2 \chi_{1-\frac{\alpha}{2}, n-1}^2}{n-1}$$

$$L.C. = \bar{S}^2$$

$$LSC = \frac{\bar{S}^2 \chi_{\frac{\alpha}{2}, n-1}^2}{n-1}.$$

5. Gráficos de control modificados

En algunos casos los límites especificados para los gráficos de control por la administración determinan una banda mucho más amplia que la banda de $6\sigma_W$ o que la banda determinada por la variabilidad natural del proceso. Entonces se puede plantear que la media del proceso se mueva en un intervalo relativamente grande sin afectar a la fracción de defectuosos del proceso. Es decir, se pueden determinar ciertos valores μ_I y μ_S (inferior y superior) para la media del proceso de forma que la fracción de disconformes del proceso siga siendo inferior a una cantidad δ . Este planteamiento equivale a probar que la media del proceso se halla en el intervalo $[\mu_I, \mu_S]$. Suponiendo que la salida del proceso se distribuye según una Normal con media μ y desviación típica σ , se considera el siguiente intervalo de variación para la media, relativo a los límites de especificación:

$$\begin{aligned} \mu_I &= LIE + Z_{\delta/2}\sigma \\ \mu_S &= LSE - Z_{\delta/2}\sigma, \end{aligned}$$

que garantizan que la fracción de disconformes del proceso se encuentra por debajo de δ , siendo Z_δ el percentil $(1 - \delta)100$ de la Normal estandarizada.

Se establecen entonces los siguientes límites de control para la media poblacional μ :

$$\begin{aligned} LIC &= \mu_I - \frac{3\sigma}{(n)^{1/2}} = LIE + Z_{\delta/2}\sigma - \frac{3\sigma}{(n)^{1/2}} \\ LSC &= \mu_S + \frac{3\sigma}{(n)^{1/2}} = LSE - Z_{\delta/2}\sigma + \frac{3\sigma}{(n)^{1/2}}. \end{aligned}$$

En el diseño de dichos límites se puede considerar que δ es la fracción de disconformes aceptable con una probabilidad $1 - \alpha$. Se tiene entonces que los límites de control vienen dados por

$$\begin{aligned} LIC &= \mu_I - \frac{Z_{\alpha/2}\sigma}{(n)^{1/2}} = LIE + Z_{\delta/2}\sigma - \frac{Z_{\alpha/2}\sigma}{(n)^{1/2}} \\ LSC &= \mu_S + \frac{Z_{\alpha/2}\sigma}{(n)^{1/2}} = LSE - Z_{\delta/2}\sigma + \frac{Z_{\alpha/2}\sigma}{(n)^{1/2}}. \end{aligned}$$

Si se requiere una fracción de disconformes del proceso γ que se debe rechazar con probabilidad $1 - \beta$. Entonces, los límites de control vendrían dados por

$$\begin{aligned} LIC &= \mu_I - \frac{Z_{\beta/2}\sigma}{(n)^{1/2}} = LIE - Z_{\gamma/2}\sigma - \frac{Z_{\beta/2}\sigma}{(n)^{1/2}} \\ LSC &= \mu_S + \frac{Z_{\beta/2}\sigma}{(n)^{1/2}} = LSE + Z_{\gamma/2}\sigma + \frac{Z_{\beta/2}\sigma}{(n)^{1/2}}. \end{aligned}$$

Para α, β, γ y δ especificados, se calcula n mediante la siguiente fórmula:

$$\begin{aligned} LSE - Z_{\delta/2}\sigma + \frac{Z_{\alpha/2}\sigma}{(n)^{1/2}} &= LSE + Z_{\gamma/2}\sigma + \frac{Z_{\beta/2}\sigma}{(n)^{1/2}} \\ n &= \frac{(Z_{\alpha/2} - Z_{\beta/2})^2}{(Z_{\delta/2} + Z_{\gamma/2})^2}. \end{aligned}$$

6. Gráfico de control de sumas acumuladas

En este diagrama se representan las sumas acumuladas de las desviaciones de las medias muestrales respecto a un valor objetivo μ_0 , es decir, para la l -ésima muestra se representa

$$D_l = \sum_{j=1}^l (\bar{X}_j - \mu_0) \quad l \in \mathbb{N}.$$

Se dice que el proceso está bajo control si hay una variación aleatoria de los valores de D_l en torno al valor objetivo μ_0 . Por el contrario, si dicha variación no es aleatoria se tiene que la sucesión de puntos definida por los valores de D_l se desplaza hacia arriba, lo que indica que hay una desviación de la media del proceso a un valor $\mu_1 > \mu_0$, o bien la sucesión de puntos se desplaza hacia abajo, lo que indica que hay una desviación de la media del proceso hacia un valor $\mu_2 < \mu_0$. Por tanto, el proceso se encuentra fuera de control.

Para determinar si se da alguna de las situaciones anteriores se diseña el diagrama en

forma de V, cuyos parámetros se definen en términos de las siguientes magnitudes:

$$\begin{aligned} d &= \left(\frac{2}{\delta^2} \right) \ln \left[\frac{1-\beta}{\alpha} \right] \\ \theta &= \arctan \left(\frac{\Delta}{2k} \right) \\ \delta &= \frac{\Delta}{\sigma_{\bar{X}}}, \end{aligned}$$

siendo Δ el cambio de la media del proceso que se quiere detectar, $k = 2\sigma_{\bar{X}}$, α denota la probabilidad de considerar que el proceso está fuera de control cuando en realidad está bajo control, β es la probabilidad de considerar que el proceso está bajo control cuando en realidad está fuera de control, θ denota el ángulo que define la amplitud de los brazos de la V y, para el último valor D_l del estadístico de sumas acumuladas, d define la distancia de dicho valor al vértice de la V. Si las sumas acumuladas determinan un punto fuera del recinto delimitado por los brazos de la V, entonces se concluye que la variación de la media del proceso está fuera de control.

7. Gráficos de control basados en medias ponderadas

7.1. Gráfico de control de la media móvil

Para diversas muestras de tamaño n extraídas en diferentes instantes de tiempo, se define la media móvil de alcance U en el instante t como sigue:

$$M_t = \frac{\bar{X}_t + \bar{X}_{t-1} + \cdots + \bar{X}_{t-U+1}}{U}.$$

Este estadístico renueva la información en cada instante de tiempo, eliminando el valor de la muestra más antigua y considerando en su lugar el valor medio muestral en el instante actual t . Puesto que

$$\text{Var}(M_t) = \frac{1}{U^2} \sum_{i=t-U+1}^t \text{Var}(\bar{X}_i) = \frac{1}{U^2} \sum_{i=t-U+1}^t \frac{\sigma^2}{n} = \frac{\sigma^2}{Un},$$

se tiene que los límites de control para el diagrama de la media móvil vienen dados por

$$\bar{\bar{X}} \pm \frac{3\sigma}{(nU)^{1/2}},$$

donde $\bar{\bar{X}}$ se calcula a partir de m muestras piloto.

Se obtiene más protección cuando se aumenta el alcance de la media móvil.

7.2. Gráfico de control de la media geométrica

En el gráfico de la media móvil los pesos o ponderaciones en cada instante de tiempo considerado eran: para los instantes de tiempo $t, \dots, t - U + 1$, la unidad, y para $s < t - U + 1$, el valor cero. En el diagrama de la media geométrica se consideran las siguientes ponderaciones:

$$Z_t = r\bar{X}_t + (1-r)Z_{t-1} = r \sum_{j=0}^{t-1} (1-r)^j \bar{X}_{t-j} + (1-r)^t Z_0,$$

siendo $Z_0 = \bar{\bar{X}}$ y $r \in (0, 1]$. El estadístico Z_t representa la media geométrica en el instante t . Este diagrama también se conoce como diagrama de la media móvil ponderada exponencialmente. Se basa en la formulación de la técnica de suavizado exponencial simple, utilizada en el análisis de series temporales. Cuando t es suficientemente grande se puede considerar la siguiente aproximación asintótica de la varianza de dicho estadístico:

$$\sigma_{Z_t}^2 \longrightarrow \sigma_Z^2 = \frac{\sigma^2}{n} \left[\frac{r}{2-r} \right].$$

Los límites de control del diagrama de la media geométrica se definen entonces como sigue:

$$\bar{\bar{X}} \pm 3\sigma \left(\frac{r}{(2-r)n} \right)^{1/2}.$$

8. Gráficos de control multivariantes

Dependiendo del proceso de fabricación y del producto, en ocasiones es necesario realizar un estudio conjunto de varias variables o características para concluir sobre la calidad y utilidad de una pieza fabricada. Estas características de la calidad pueden ser independientes o estar relacionadas, pero en ambos casos las probabilidades de considerar que el proceso está fuera de control cuando en realidad está bajo control, es decir, la probabilidad de error de tipo I, y de representar correctamente un punto bajo control cambian dependiendo de que se considere un diagrama individual para la media de cada característica, o bien un diagrama conjunto para las medias de todas las características. Por ejemplo, si se consideran dos características X_1 y X_2 , variables independientes, entonces la probabilidad de exceder los límites de control en el diagrama conjunto de ambas variables es considerablemente mayor que las probabilidades de exceder los límites de control en los diagramas individuales. Por tanto, el estudio individual de ambas variables distorsiona bastante el control conjunto de éstas. En general, si $X_1, \dots, X_p$, son características independientes que se desea controlar conjuntamente, entonces

$$P[\text{error de tipo I del diagrama conjunto}] = 1 - (1 - \alpha)^p,$$

es decir,

$$P[\text{concluir que el proceso está bajo control cuando en realidad lo está}] = (1 - \alpha)^p,$$

donde se está denotando por α la probabilidad de error de tipo I para cada característica, es decir, en los diagramas individuales. En el caso en que dichas variables no son independientes es difícil medir la distorsión ocasionada por un estudio individualizado de las mismas. Seguidamente se estudia el problema de control de varias variables relacionadas, es decir, el problema de control multivariante, a partir de gráficos de control diseñados en términos de un estadístico con distribución $\mathcal{T}^2$ de Hötelling.

Se estudia en primer lugar el caso bivalente. Supongamos que se quiere realizar un control conjunto de dos características de la calidad cuya distribución es una Normal bivalente con vector de medias (μ_1, μ_2) y con matriz de covarianzas Σ . A partir de muestras preliminares, se estima el vector de medias en términos de $(\bar{X}_1, \bar{X}_2)$ y la matriz de covarianzas mediante la matriz de cuasivarianzas y cuasicovarianzas muestrales. Se tiene entonces que el estadístico

$$T^2 = \frac{n}{S_1^2 S_2^2 - S_{12}^2} \left[S_2^2 (\bar{X}_1 - \bar{\bar{X}}_1)^2 + S_1^2 (\bar{X}_2 - \bar{\bar{X}}_2)^2 - 2S_{12}(\bar{X}_1 - \bar{\bar{X}}_1)(\bar{X}_2 - \bar{\bar{X}}_2) \right] \quad (1)$$

se distribuye según una $\mathcal{T}^2$ de Hötelling con 2 y $n - 1$ grados de libertad, siendo n el tamaño muestral, $\bar{X}_i$, $i = 1, 2$, las medias muestrales, S_i^2 , $i = 1, 2$, las cuasivarianzas muestrales y S_{ij} , $i, j = 1, 2$, $i \neq j$, las cuasicovarianzas para las dos características de la calidad estudiadas, evaluadas en la muestra considera.

Metodología. Si $T^2 > \mathcal{T}_{\alpha, 2, n-1}^2$, entonces al menos una de las características estudiadas se encuentra fuera de control, rechazándose la hipótesis de que el proceso está bajo control, siendo $\mathcal{T}_{\alpha, 2, n-1}^2$ el punto de la distribución $\mathcal{T}^2(2, n - 1)$ que deja a la derecha una probabilidad α . Gráficamente, si $S_{12} = 0$ y $\bar{X}_1$ y $\bar{X}_2$ son independientes, entonces la ecuación (1) define una elipse con los ejes principales paralelos a los ejes de abscisas y ordenadas donde se representan, respectivamente, los valores de $\bar{X}_1$ y $\bar{X}_2$. Este diagrama tipo elipse para el control bivalente de dos características presenta algunos inconvenientes:

- Se pierde la secuencia en el tiempo (evolución del proceso) y, por ejemplo, no se pueden apreciar patrones de no aleatoriedad tales como las rachas.
- La representación gráfica de más de dos características se complica.

Los puntos anteriores motivan el uso de gráficos de control, donde los valores del estadístico (1) se representan en un diagrama cartesiano y el valor $\mathcal{T}_{\alpha, 2, n-1}^2$ se representa mediante una línea horizontal, paralela al eje de abscisas, que define el único límite superior de control. Para cada muestra, se evalúa el estadístico (1): si el punto correspondiente a dicho valor supera el límite superior de control, entonces el proceso está fuera de control; en caso contrario, el proceso está bajo control.

La generalización del planteamiento anterior al caso de p variables sería como sigue.

Se considera el vector aleatorio $(X_1, \dots, X_p)$ que representa las p características de la calidad que se pretende analizar conjuntamente. Análogamente al caso bivalente, para una muestra de tamaño n se calcula el estadístico

$$T^2 = n(\bar{x} - \bar{\bar{x}})S^{-1}(\bar{x} - \bar{\bar{x}}),$$

donde el vector de medias muestrales $\bar{x}$ se evalúa en la muestra considerada, así como la matriz de cuasivarianzas y cuasicovarianzas muestrales S . El vector $\bar{\bar{x}}$ se calcula a partir de un conjunto de muestras preliminares. El estadístico T^2 sigue una distribución T^2 de Hötelling con p y $n - 1$ grados de libertad. El diagrama se define en términos del límite superior de control $\mathcal{T}_{\alpha,p,n-1}^2$. Se suele utilizar la siguiente aproximación:

$$\mathcal{T}_{\alpha,p,n-1}^2 = \frac{p(n-1)}{n-p} \mathcal{F}_{\alpha,p,n-p},$$

siendo $\mathcal{F}_{\alpha,p,n-p}$ el valor de la distribución F de Snedecor que deja a la derecha un área α y considerando $n > p$.

A partir de m muestras piloto, los cálculos para realizar las estimaciones necesarias para evaluar T^2 serían los siguientes. Denotando mediante X_{ijk} la variable aleatoria que representa el valor en la unidad i de la característica j en la k -ésima muestra, se tienen los siguientes estadísticos muestrales, para $j = 1, \dots, p$, y $k = 1, \dots, m$:

$$\begin{aligned} \bar{X}_{jk} &= \frac{1}{n} \sum_{i=1}^n X_{ijk} \\ S_{jk}^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_{ijk} - \bar{X}_{jk})^2 \\ S_{jhk} &= \frac{1}{n-1} \sum_{i=1}^n (X_{ijk} - \bar{X}_{jk})(X_{ihk} - \bar{X}_{hk}), \quad j \neq h \\ \bar{\bar{X}}_j &= \frac{1}{m} \sum_{k=1}^m \bar{X}_{jk} \\ \bar{\bar{S}}_j^2 &= \frac{1}{m} \sum_{k=1}^m S_{jk}^2 \\ \bar{\bar{S}}_{jh} &= \frac{1}{m} \sum_{k=1}^m S_{jhk}. \end{aligned}$$