

1. TEMA 4: EL CONTROL DE FABRICACIÓN POR VARIABLES

Contenido del tema: Introducción. Gráficos de control de la media. Capacidad del proceso. Curvas OC. El ARL en el control por variables. Ejercicios del tema.

1.1. Introducción.

Cuando la característica de calidad puede expresarse en términos de una variable aleatoria continua (longitud, peso, resistencia,...) tendrá interés controlar el valor medio y la variabilidad, para ello existe toda una teoría que se engloba dentro del denominado control por variables.

Supondremos que queremos controlar una característica medible, X , cuyo objetivo o valor nominal, (en inglés Target: blanco u objetivo) se representará por T . Así, si X representa a la variable aleatoria que mide la longitud de una pieza, T será la medida especificada por el diseño de la pieza en cuestión. En este caso, supondremos que el proceso se ajusta de manera que la media de la variable aleatoria X es precisamente T .

1.2. Gráficos de control para variables

Para determinar la capacidad de un proceso debemos obtener datos de su funcionamiento en condiciones normales, es decir tratando de eliminar todas las causas asignables de variación. En ese supuesto, estimar la capacidad equivale a estimar la desviación típica de la característica medible, X .

Dado que durante la propia recogida de información podría ocurrir que el proceso pasase a una situación fuera de control, es aconsejable tomar varias muestras de tamaño pequeño e igualmente espaciadas a lo largo del intervalo de producción. Denótese por x_{ij} el j -ésimo elemento de la i -ésima muestra y considérese k , el número de muestras, y n , el tamaño de las mismas. Si el proceso hubiese permanecido bajo control a lo largo de todo el período de recogida de información, puede estimarse la varianza mediante la cuasivarianza muestral conjunta

$$\hat{s}^2 = \frac{\sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{\bar{x}})^2}{nk - 1},$$

siendo $\bar{\bar{x}}$ la media muestral conjunta de los nk datos.

$$\bar{\bar{x}} = \frac{\sum_i \bar{x}_i}{k} = \frac{\sum_i \sum_j x_{ij}}{N}, \text{ con } N = kn,$$

No debe olvidarse que el estado de control no es algo espontáneo sino un logro en sí mismo, por ello es probable que no todas las k muestras provengan de la misma población. Por este motivo se utilizan los gráficos de control.

El objetivo del control de procesos es detectar, lo antes posible, que el proceso está fuera de control. Una vez detectada la salida de control del proceso debe iniciarse una investigación con el objeto de encontrar cual es la causa asignable, eliminarla del proceso y volver a la situación bajo control. Este método asegura que la variabilidad final que recibe el cliente sea mucho menor que si no se realizara el control, independientemente de que todas las piezas puedan seguir siendo correctas en cuanto a que pertenecen al intervalo de tolerancias.

Para el diseño de este tipo de gráficos es necesario especificar el tamaño de la muestra, la amplitud de los límites de control y el intervalo de muestreo que se va a utilizar. Para dar una solución exacta a este problema de estimación es necesario contar con la ayuda de las personas involucradas en el proceso y tener también en cuenta los factores económicos del muestreo. En un primer estudio analizaremos la estimación de este tipo de gráficos dando unas directrices genéricas sin valorar estos aspectos económicos.

Para la construcción de los gráficos de control supondremos que se toman de k muestras aleatorias de tamaño n de los objetos de interés que se suponen homogéneas. A continuación se siguen los siguientes pasos que serán diferentes según se estén utilizando las desviaciones típicas muestrales s_i o los recorridos R_i .

1.2.1. Gráficos de control (\bar{x}, R)

Para el control de una característica medible y representada por una variable aleatoria X , existen dos tipos de gráficos, por una parte gráficos que nos ayuden a controlar la posición (gráficos de la media o mediana) y por otra los que controlan la dispersión (gráficos de desviaciones, varianzas o recorridos). Existiendo distintos gráficos para el control de variables pero, por tradición, uno de los más populares son los basados en rangos, básicamente por su gran facilidad de cálculo.

El rango de una muestra no es más que la diferencia entre la mayor y la menor de sus observaciones. Bajo la hipótesis de que la distribución es normal, puede calcularse la distribución del rango muestral. En particular, su media es $d_2\sigma$, donde d_2 es una constante que depende del tamaño muestral, n , y que está tabulada. También están tabulados los coeficientes D_1 y D_2 que definen el intervalo en el que se debe encontrar el rango muestral con una probabilidad del 99% para una población normal de varianza 1.

En procesos bajo control, en los que se supone σ conocida, puede construirse un gráfico de control para la dispersión, medida ésta mediante rangos. Tomando como línea central el valor esperado del rango muestral ($d_2\sigma$) y como límites de control los dados por las constantes tabuladas ($D_1\sigma$ y $D_2\sigma$), se realiza una gráfica en la que se representan los rangos muestrales. Cada vez que uno de ellos se salga de los límites de control diremos que el proceso no está bajo control.

Puesto que R_i/d_2 es un estimador insesgado de σ y, por tanto, su media (a lo largo de todas las muestras) también lo será. Definiendo

$$\bar{R} = \frac{\sum_i R_i}{k},$$

se tiene entonces que podemos estimar σ mediante \bar{R}/d_2 y construir los límites de control de los gráficos de medias (llamados ahora gráficos (\bar{x}, R)) mediante

$$\bar{\bar{x}} \pm \frac{3\bar{R}}{\sqrt{nd_2}}.$$

En donde $\bar{\bar{x}}$ es la media global de todos los datos o, lo que es equivalente, la media de las medias de cada submuestra:

$$\bar{\bar{x}} = \frac{\sum_i \bar{x}_i}{k} = \frac{\sum_i \sum_j x_{ij}}{N}, \text{ con } N = kn,$$

Si todas las muestras provienen de la misma distribución entonces se tiene que

$$\frac{\sqrt{nd_2}(\bar{x}_i - \bar{\bar{x}})}{\bar{R}}$$

tiene distribución aproximada $N(0, 1)$. Con lo cual cada valor \bar{x}_i debe estar en el intervalo

$$\left[\bar{\bar{x}} - \frac{3\bar{R}}{\sqrt{nd_2}}, \bar{\bar{x}} + \frac{3\bar{R}}{\sqrt{nd_2}} \right] = [\bar{\bar{x}} - A_2\bar{R}, \bar{\bar{x}} + A_2\bar{R}]$$

con probabilidad aproximada del 99,73. Utilizando intervalos de amplitud 3 sigma se tiene una probabilidad de cometer un falsa alarma o error de primera especie α igual al 2,7 por mil, ya que $1 - \alpha = 0,0027$. Las constantes A_2 están tabuladas para distintos valores de n .

Podremos obtener los valores de los límites de control 3σ para la media como:

$$\begin{aligned} LCL &= \bar{\bar{x}} - A_2\bar{R} \\ CL &= \bar{\bar{x}} \\ UCL &= \bar{\bar{x}} + A_2\bar{R} \end{aligned}$$

Análogamente, el gráfico de variabilidad mediante rangos tiene una línea central con altura \bar{R} y líneas de control $D_3\bar{R}$ y $D_4\bar{R}$, estando D_3 y D_4 también tabuladas.

Para la determinación de D_3 y D_4 se debe de tener en cuenta que la idea es construir límites de control inferior y superior para el recorrido medio (la línea central en este caso será \bar{R}). Si suponemos que $\sigma_R = d_3\sigma$, un estimador para σ_R será:

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

por lo que podremos obtener como límites de control 3σ :

$$\begin{aligned} LCL &= \bar{R} - 3\hat{\sigma}_R = \bar{R} - 3d_3 \frac{\bar{R}}{d_2} = D_3\bar{R} \\ CL &= \bar{R} \\ UCL &= \bar{R} + 3\hat{\sigma}_R = \bar{R} + 3d_3 \frac{\bar{R}}{d_2} = D_4\bar{R} \end{aligned}$$

donde las constantes, antes mencionadas D_3 y D_4 , que vienen dadas por:

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

están tabuladas en función de los valores de n .

1.2.2. Gráficos de control (\bar{x}, s)

Como se comentó anteriormente para el caso de los recorridos, partiremos de k muestras de tamaño n . En este caso los pasos para la construcción del gráfico (\bar{x}, s) serán los siguientes:

1. Calcular la media y la desviación típica de cada muestra, \bar{x}_i y s_i , respectivamente ($i = 1, 2, \dots, k$).
2. Calcular la media y la desviación típica global. La media global no es más que la media de las medias muestrales, que puede calcularse promediando las medias de cada submuestra o bien promediando el total de los datos (kn)

$$\bar{\bar{x}} = \frac{\sum_i \bar{x}_i}{k} = \frac{\sum_i \sum_j x_{ij}}{N}, \text{ con } N = kn,$$

Esta media resulta ser un estimador insesgado de la media teórica, μ (también denominada por T en este contexto). Como, la desviación típica muestral es un estimador sesgado (aunque asintóticamente insesgado) de la desviación típica teórica. De hecho se tiene $E[s_i] = c_2\sigma$, donde c_2 es una constante, dependiente de n , que suele venir tabulada. De esta forma, s_i/c_2 es un estimador insesgado de σ y ello motiva el utilizar

$$\hat{\sigma} = \frac{\bar{s}}{c_2} = \frac{\sum_i s_i}{c_2 k}$$

como estimador (centrado) de la desviación típica teórica.

Si la varianza fuese la misma en todas las muestras, el estimador propuesto es ligeramente peor que el basado en la media de las cuasivarianzas

$$\hat{s}_T^2 = \frac{\sum_i \hat{s}_i^2}{k} = \frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)^2}{k(n-1)} = \frac{n}{N-k} \sum_i s_i^2,$$

que es insesgado para σ^2 , pero cuya raíz cuadrada no lo es para σ . La razón de usar $\hat{\sigma} = \bar{s}/c_2$ es que es más robusto.

3. Contrastar si todas las medias son homogéneas. Si todas las muestras provienen de la misma distribución entonces se tiene que

$$\frac{\sqrt{nc_2}(\bar{x}_i - \bar{\bar{x}})}{\bar{s}}$$

tiene distribución aproximada $N(0, 1)$. Con lo cual cada valor \bar{x}_i debe estar en el intervalo

$$\left[\bar{\bar{x}} - \frac{3\bar{s}}{\sqrt{nc_2}}, \bar{\bar{x}} + \frac{3\bar{s}}{\sqrt{nc_2}} \right]$$

con probabilidad aproximada del 99,73. Utilizando intervalos de amplitud 3 sigma se tiene una probabilidad de cometer un falsa alarma o error de primera especie α igual al 2,7 por mil, ya que $1 - \alpha = 0,0027$.

4. Podremos obtener los valores de los límites de control 3σ para la media como:

$$\begin{aligned} LCL &= \bar{\bar{x}} - A_1 \bar{s} \\ CL &= \bar{\bar{x}} \\ UCL &= \bar{\bar{x}} + A_1 \bar{s} \end{aligned}$$

en donde $A_1 = \frac{3}{\sqrt{nc_2}}$.

Para ejecutar el contraste de que las k muestras, lo que equivale a realizar k contrastes de hipótesis, se suele hacer uso de un gráfico en el que se representa en abscisas el tiempo (el numero de muestras elegidas de forma temporal) y

en ordenadas la media de la muestra en cuestión. Dibujando dos líneas horizontales, o límites de control, de altura igual a los extremos del intervalo anteriormente mencionado, se comprueba fácilmente cuando alguna de las medias en cuestión cae fuera de los límites de control, siendo esto indicación de que el proceso, en ese instante, está fuera de control.

1.2.3. Gráfico de desviaciones típicas s

Este es un gráfico en el que se mide la variabilidad de cada muestra haciendo uso de su desviación típica, s_i . Puede demostrarse que un intervalo que contiene a la desviación típica muestral s_i con probabilidad aproximada de un 99,73 % es el dado por $(B_3\bar{s}, B_4\bar{s})$ donde las constantes B_3 y B_4 (que dependen de n) están tabuladas.

Para construir este gráfico de control se procede de la siguiente forma:

1. Dado n , encontrar los valores B_3 y B_4 y calcular \bar{s} . Con todo ello obtener los límites inferior $B_3\bar{s}$ y superior $B_4\bar{s}$ de control.
2. Construir un gráfico que presente en abscisas el tiempo y en ordenadas los valores s_i . Trazar también las líneas central (\bar{s}) y los límites de control ($B_3\bar{s}$ y $B_4\bar{s}$).
3. Si alguna de las desviaciones típicas se sale de los límites de control, se admite que el proceso no está bajo control en ese instante.

1.2.4. Gráficos de control (\bar{x}, s)

Es frecuente, la mayoría de los programas de ordenador lo hacen, usar las cuasidesviaciones típicas, \hat{s}_i , en lugar de las desviaciones típicas muestrales, en este caso los valores de las constantes cambian. De hecho se tiene $E[\hat{s}_i] = c_4\sigma$, donde c_4 es una constante, dependiente de n , que suele venir tabulada. De esta forma, \hat{s}_i/c_4 es un estimador insesgado de σ y ello motiva usar:

$$\hat{\sigma} = \frac{\bar{\hat{s}}}{c_4} = \frac{\sum_i \hat{s}_i}{c_4 k}$$

como estimador (centrado) de la desviación típica teórica.

El valor c_4 viene dado por la expresión:

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

Podremos obtener los valores de los límites de control 3σ para la media como:

$$\begin{aligned}
LCL &= \bar{\bar{x}} - A_3 \bar{\hat{s}} \\
CL &= \bar{\bar{x}} \\
UCL &= \bar{\bar{x}} + A_3 \bar{\hat{s}}
\end{aligned}$$

en donde $A_3 = \frac{3}{\sqrt{nc_4}}$.

En este supuesto que $\bar{\hat{s}}/c_4$ se utiliza para estimar la desviación típica, se pueden obtener los siguientes límites para el gráfico de dispersión:

$$\begin{aligned}
LCL &= B_3 \bar{\hat{s}} \\
CL &= \bar{\hat{s}} \\
UCL &= B_4 \bar{\hat{s}}
\end{aligned}$$

en donde $B_3 = 1 - \frac{3}{c_4} \sqrt{1 - c_4^2}$ y $B_4 = 1 + \frac{3}{c_4} \sqrt{1 - c_4^2}$.

1.2.5. Otros gráficos de control para variables

Aunque los gráficos vistos anteriormente son los más utilizados, existen otras opciones.

Gráficos con valores poblacionales conocidos

Cuando se conocen la media o la desviación de la variable a estimar podría utilizarse estos valores para la construcción de los diagramas de control:

$$\begin{aligned}
LCL &= \mu - A\sigma \\
CL &= \mu \\
UCL &= \mu + A\sigma
\end{aligned}$$

en donde $A = \frac{3}{\sqrt{n}}$.

Del mismo modo para la construcción de un gráfico tipo R conociendo la desviación σ , teniendo en cuenta que $\sigma = R/d_2$, y que $\sigma_R = d_3\sigma$ quedaría:

$$\begin{aligned}
LCL &= d_2\sigma - d_3\sigma = D_1\sigma \\
CL &= d_2\sigma \\
UCL &= d_2\sigma + d_3\sigma = D_2\sigma
\end{aligned}$$

en donde $D_1 = d_2 - 3d_3$ y $D_2 = d_2 + 3d_3$.

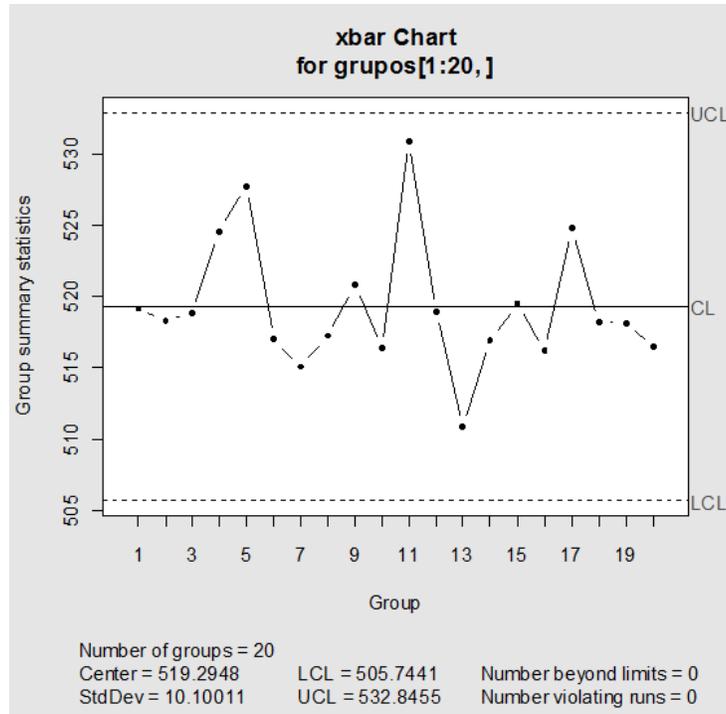


Figura 1-1 Gráfico para la media con 20 subgrupos utilizando las desviaciones típicas.

Gráficos probabilísticos

También existe la opción de utilizar gráficos basados directamente en la varianza muestral del proceso (o en la cuasivarianza) S^2 :

$$LCL = \frac{\overline{S^2}}{n-1} \chi_{\alpha/2, n-1}^2$$

$$CL = \overline{S^2}$$

$$UCL = \frac{\overline{S^2}}{n-1} \chi_{1-\alpha/2, n-1}^2$$

en donde $\chi_{\alpha/2, n-1}^2$ y $\chi_{1-\alpha/2, n-1}^2$ denotan los cuantiles correspondientes de la distribución Chi-cuadrado de Pearson con n-1 grados de libertad. En el caso que se tuviese un valor conocido de σ también podría utilizarse en la expresión anterior.

Gráficos para medidas individuales

En el supuesto que las submuestras utilizadas sean de un solo valor, es decir el tamaño muestral n es 1, este es un caso frecuente en control de calidad de industrias químicas, en las que no tiene mucho sentido tomar muestras de varias unidades de la misma mezcla (piensese en una cervecera). En este suelen usarse medias móviles de distinto orden o la amplitud móvil entre dos observaciones sucesivas para estimar la variabilidad del proceso, a partir de estas amplitudes

se estimaría la media de esos rangos. Así si se utilizasen sólo dos valores las constantes utilizadas para el gráfico serían las equivalentes a las usadas de una muestra de tamaño $n = 2$ que resulta un valor de $d_2 = 1,128$.

Estimación de la capacidad

Si en alguno de los dos gráficos anteriormente citados hay alguna muestra que se sale de los límites de control, la eliminaremos, pasando a representar nuevamente los gráficos excluyendo los valores fuera de control. Si en ese momento todas las muestras se hallan bajo control, las consideraremos un grupo homogéneo y en caso contrario procederemos como se acaba de indicar, eliminando las muestras fuera de control y repitiendo el proceso hasta obtener un grupo homogéneo.

Seguidamente se contrasta la normalidad de la distribución de fabricación con los datos de los casos que han quedado bajo control. Para ello pueden utilizarse contrastes como el chi-cuadrado, Kolmogorov-Smirnov-Lilliefors, o el de Shapiro-Wilk (entre otros).

Una vez comprobada la normalidad, la varianza se estimará mediante \bar{s}/c_2 . En procesos mal controlados (es decir, con muchas muestras heterogéneas) el estimador anterior puede ser poco preciso. En tal caso, es conveniente tomar la estimación como un valor inicial en base al cual comenzar a controlar el proceso.

1.2.6. El concepto de capacidad

En todo proceso productivo se puede definir una medida de la capacidad que el proceso tiene para cumplir para satisfacer sus especificaciones de calidad. Es el llamado *índice de capacidad*.

Si la característica de calidad es continua y con distribución $N(T, \sigma)$, cuando el proceso se encuentra bajo control casi la totalidad de las unidades producidas (exáctamente el 99,73 % de las mismas) se encuentran en un intervalo de amplitud 6σ , centrado en el valor nominal, T . Matemáticamente,

$$\begin{aligned} P(T - 3\sigma \leq X \leq T + 3\sigma) &= P(|X - T| \leq 3\sigma) = \Phi(3) - \Phi(-3) = \\ &2\Phi(3) - 1 = 2 \cdot 0,998650 - 1 = 0,9973 \end{aligned}$$

Por ello, el intervalo $[T - 3\sigma, T + 3\sigma]$ se denomina de tolerancias naturales o intrínsecas del proceso. Teniendo esto en cuenta se define la *capacidad del proceso como 6σ* .

Para analizar la adecuación de un proceso a las especificaciones preestablecidas es absolutamente necesario conocer su capacidad. Supóngase que el intervalo de tolerancias de fabricación de un proceso es (LS, LI) , centrado en T .

Entonces, se define el índice de capacidad del proceso como

$$IC = \frac{LS - LI}{6\sigma}.$$

Atendiendo al valor del índice de capacidad se presentan tres casos diferenciados:

1. $IC < 1$. El proceso no puede cumplir las especificaciones requeridas. Además, cuanto más pequeño sea el índice de capacidad mayor será la proporción de piezas defectuosas. La recomendación práctica será tratar de reducir la variabilidad no asignable, lo cual requiere cambios en el proceso o en el producto. Si esto no es posible, el proceso ha de someterse a una inspección muy frecuente para prevenir cualquier pequeño desajuste que aumente, aún más, el número de defectuosos. Nunca deben modificarse las tolerancias, que deben estar basadas en especificaciones técnicas y razones económicas, pero no en la capacidad del proceso.
2. $IC = 1$. Según lo apuntado anteriormente, el proceso fabricará aproximadamente un 0,3 % de defectuosos. Tradicionalmente un tipo de proceso que cumpliera esta característica se consideraba justamente apto para la fabricación. No obstante, en la actualidad este porcentaje de defectuosos puede ser inaceptable en muchos productos (cuyas proporciones de defectuosos se cuentan en uno o dos por millón). Además también exige un control muy estricto porque pequeños desplazamientos de la media pueden provocar aumentos muy considerables de la proporción de defectuosos.
3. $IC > 1$. En este caso el porcentaje de defectuosos es realmente muy pequeño (aunque la palabra “pequeño” siempre es relativa ya que depende del sector industrial concreto). Por este motivo tan sólo es necesario supervisar el proceso por si se producen desviaciones acusadas del estado de control.

El índice de capacidad nos permite calcular el porcentaje de defectuosos, pero además también ayuda a elegir entre procesos alternativos, a establecer un sistema de control durante la fabricación mediante la toma de muestras cada cierto tiempo (ver tabla adjunta) e indica un punto de partida para la mejora del proceso. En la metodología Six Sigma este índice se busca que de un número superior a 2.

Índice de capacidad	Frecuencia de inspección
$IC \leq 1$	Todas las unidades
$1 < IC \leq 1,4$	Intensiva (cada 15 ó 30 minutos)
$1,4 < IC \leq 1,7$	Moderada (cada hora)
$1,7 < IC \leq 2$	Espaciada (cada 2 horas)
$2 < IC$	Depende de la frecuencia de anomalías

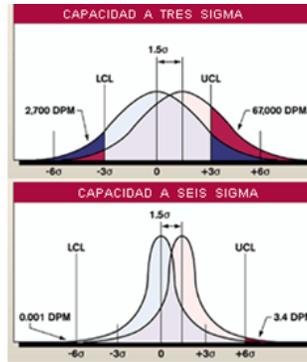


Figura 1-2 Concepto de capacidad según seis sigma.

Example 1 *La desviación típica de un proceso de fabricación de componentes electrónicos es 0,05 V. Se trata de fabricar dos componentes A y B que se consideran defectuosos si su tensión varía más de 0,5 V. Calcular el índice de capacidad del proceso si el coste de fabricación de ambos es el mismo, 50 euros. pero el coste de fallo para el usuario es de 80 euros para A y 250 para B.*

Suponiendo una función de coste cuadrática, los costes sociales en cada caso son

$$C_A(x) = \frac{80}{0,25}(x - T)^2, \quad C_B(x) = \frac{250}{0,25}(x - T)^2,$$

luego, los intervalos de tolerancias son $[T - 0,395, T + 0,395]$, para el A, y $[T - 0,224, T + 0,224]$, para el B. Como consecuencia, sus índices de capacidad son $IC_A = \frac{0,79}{0,30} = 2.63$ y $IC_B = \frac{0,448}{0,30} = 1.49$. Así pues el proceso A es muy capaz y la inspección es poco frecuente, mientras que en el B, la capacidad es adecuada y la inspección debe ser moderada. Desde la perspectiva Six Sigma el proceso B no sería capaz, mientras que el A sí.

Example 2 *Para determinar la capacidad de un proceso se toman $k = 25$ muestras de tamaño $n = 6$ y se calculan la media y la desviación típica de cada una de ellas. A continuación se adjuntan los resultados de las medias de cada muestra, indicando entre paréntesis la desviación típica: 53 (9,3), 51 (3,1), 48 (2,5), 48 (1,3), 48 (6,7), 51 (7,9), 49 (3,8), 51 (8,2), 48 (2,6), 49 (5,4), 47 (4,4), 47 (5,1), 57 (2,9), 51 (4,1), 48 (1,4), 49 (4,2), 52 (5,3), 50 (6,6), 51 (6,7), 51 (8,1), 49 (6,0), 50 (7,9), 49 (9,8), 46 (4,2), 47 (3,5).*

Pueden calcularse fácilmente los límites para los gráficos de control. Para el gráfico de medias las líneas de control son $\bar{\bar{x}} + \frac{3\bar{s}}{\sqrt{nc_2}} = 56,34$ y $\bar{\bar{x}} - \frac{3\bar{s}}{\sqrt{nc_2}} = 42,86$, con línea central en 49,6. Por su parte, para el gráfico s, las líneas de control son $B_4\bar{s} = 10,32$ y $B_3\bar{s} = 0,1597$, con línea central en 5,24. A partir del total de 25 muestras la estimación para la media del proceso es 49,6 y para la desviación típica, 5,51.

Mediante los gráficos de control (\bar{x}, s) y s , resulta fácil ver que la muestra número 13 (con una media de 57 y una desviación típica de 2,9) se sale de

control en el gráfico de medias. Si se procede a eliminarla y se calculan nuevamente los límites de control (56,16 y 42,42, para el de medias, con línea central en 49,29, y 10,51 y 0,1627, para el de desviaciones típicas, con l ínea central en 5,34)

Ahora ya no hay ninguna muestra que se salga de control y, por tanto, finaliza el procedimiento para determinar la estimación de la media y la varianza del proceso, siendo esos valores $\bar{x} = 49,29$ y $\hat{\sigma} = 5,61$. De hecho, el intervalo de tolerancias naturales del proceso sería (32.46, 66.12), de longitud 33.66.

Example 3 Determinar el número de submuestras k , en un gráfico de control para la media, en el que queremos elegir submuestras de tamaño 4, y pretendemos que la diferencia entre el valor nominal y la media muestral no sea superior a 5 milésimas con un riesgo $\alpha = 0,1$, suponiendo una desviación típica de 0.02.

Utilizando la fórmula para un intervalo de confianza en el que el tamaño muestral sería kn , se obtiene

$$\begin{aligned} kn &= \left(\frac{\sigma \cdot Z_{\alpha/2}}{\varepsilon} \right)^2 \\ k &= \frac{1}{4} \left(\frac{0,02(2,575)}{0,005} \right)^2 \\ k &= 26,523 \end{aligned}$$

Por tanto habría que tomar un total de 27 submuestras de tamaño 4.

Una vez determinada la capacidad de un proceso, se compara con las tolerancias del producto, para hallar así el índice de capacidad y seleccionar la frecuencia de muestreo. Así se da paso al llamado control de fabricación.

El objetivo del control estadístico de fabricación es comprobar continuamente si el proceso se encuentra bajo control. Es un error muy frecuente el pensar que el control de fabricación se introduce para observar si el producto se halla dentro de tolerancias (cambiando los l í mites de control por los de tolerancia).

Construyendo gráficos de medias y de variabilidad puede observarse si, a lo largo del tiempo, el proceso está bajo control, simplemente viendo que en ninguno de los dos gráficos las observaciones se escapan de los l í mites de control. Si, por el contrario, en alguno de estos grá ficos una muestra se sale de control, dado que la probabilidad de error de tipo I (afirmar que la muestra no está bajo control cuando realmente lo está) es muy pequeña (del orden del 1), rechazaremos el que el proceso se encuentre bajo control en ese instante e investigaremos las causas de este hecho para evitarlas en el futuro.

1.3. Curva característica de operación (Curva OC)

Con objeto de valorar la idoneidad de los gráficos de control para detectar cambios en el proceso se puede recurrir a la teoría de los contrastes de hipótesis, en donde la hipótesis nula sería considerar que el proceso está bajo control frente a la alternativa que sería que el proceso no está bajo control, bajo estas premisas surge los conceptos de error de tipo I que sería el error que se comete al suponer el proceso fuera de control cuando realmente está bajo control, también llamado en este contexto *riesgo del vendedor o proveedor* o *riesgo α* (también se puede llamar probabilidad de falsa alarma) y el error de tipo II; que es el que se comete al considerar que el proceso si está bajo control no estándolo, llamado riesgo β o *riesgo del comprador*. Estos riesgos dependen del tamaño de la muestra y de los intervalos de aceptación y rechazo de la hipótesis nula. Cuando se grafican el riesgo β en función de la magnitud del cambio que se pretende estudiar (normalmente expresado en unidades de desviación estandar) se obtiene una curva denominada *Curva Característica de Operación* o *curva OC* (Operating Characteristics, que sería la equivalente a la curva de potencia de un contraste, concretamente sería uno menos la potencia).

En la figura 1-3 se muestra la curva OC para el control de la media del proceso, mientras que la figura 1-4 muestra como varían estas curvas en función del tamaño muestral de cada submuestra, gráficamente la probabilidad de aceptar en función de valores de la media en el primer caso y en función de la desviación típica en el segundo caso.

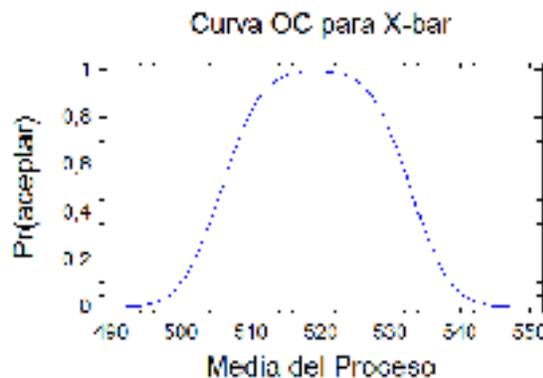


Figura 1-3 Curva OC para la media del proceso.

En el caso de que estemos interesados en calcular el error de tipo II al pasar de un valor nominal μ a un valor $\mu_1 = \mu_0 + k\sigma$, supuesto que la característica de calidad medible se ajusta a una distribución $N(\mu, \sigma)$, con lo que la media muestral $\bar{x} \in N(\mu, \frac{\sigma}{\sqrt{n}})$, si suponemos que $UCL = \mu_0 + d\frac{\sigma}{\sqrt{n}}$ y $LCL = \mu_0 - d\frac{\sigma}{\sqrt{n}}$

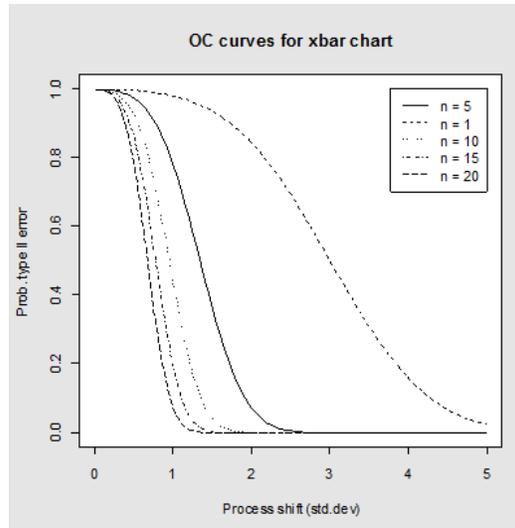


Figura 1-4 Diferentes curvas OC en función del tamaño de las submuestras.

$$\begin{aligned} \beta &= P(LCL \leq \bar{x} \leq UCL / \mu = \mu_1 = \mu_0 + k\sigma) \\ \beta &= \Phi \left[\frac{UCL - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] - \Phi \left[\frac{LCL - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] \\ \beta &= \Phi \left[\frac{\mu_0 + d\frac{\sigma}{\sqrt{n}} - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] - \Phi \left[\frac{\mu_0 - d\frac{\sigma}{\sqrt{n}} - (\mu_0 + k\sigma)}{\frac{\sigma}{\sqrt{n}}} \right] \\ \beta &= \Phi [d - k\sqrt{n}] - \Phi [-d - k\sqrt{n}] \end{aligned}$$

En el supuesto más clásico en que se usen límites 3 sigma ($d = 3$), tomando por ejemplo $n = 5$ y $k = 1$, es decir $\mu_1 = \mu_0 + \sigma$, tendríamos:

$$\begin{aligned} \beta &= \Phi [3 - 1\sqrt{5}] - \Phi [-3 - 1\sqrt{5}] \\ &= \Phi [0,76393] - \Phi [-5,2361] \\ &\simeq 0,7764 \end{aligned}$$

En este supuesto el error de tipo II tiene una alta probabilidad, si lo que se pretende es medir el riesgo cuando $\mu_1 = \mu_0 + 2\sigma$, realizando cálculos similares se obtendría un valor de 0,07 para la probabilidad de error de tipo II.

Para la construcción de la curva OC para un gráfico de variables se puede hacer una gráfica en la que en el eje Y se ponga la probabilidad del error de tipo II o riesgo β y en el eje de abscisas distintos valores para k , según la expresión $\beta = \Phi [d - k\sqrt{n}] - \Phi [-d - k\sqrt{n}]$, de esta forma es posible obtener varias curvas OC para cada uno de los tamaños de las submuestras consideradas n .

Dados los puntos en la curva característica $(0, \alpha)$ y (μ_1, β) teniendo en cuenta la teoría de contraste de hipótesis, suponiendo una que los datos siguen una distribución normal, el tamaño muestral vendría determinado por:

$$n = \left(\frac{Z_{\frac{\alpha}{2}} + Z_{\beta}}{d} \right)^2$$

siendo d el descentrado relativo, es decir,

$$d = \frac{|\mu - \mu_0|}{\sigma_0}$$

En este contexto la probabilidad de rechazar $H_0(\mu = \mu_0)$ vs $H_0(\mu \neq \mu_0)$:

$$2 - [\Phi(Z_{\frac{\alpha}{2}} - d\sqrt{n}) + \Phi(Z_{\frac{\alpha}{2}} + d\sqrt{n})]$$

La curva característica o probabilidad de aceptar H_0 en función del descentrado d , sería:

$$OC(d) = \Phi(Z_{\frac{\alpha}{2}} - d\sqrt{n}) + \Phi(Z_{\frac{\alpha}{2}} + d\sqrt{n}) - 1$$

En el supuesto de que estemos interesados en contrastar hipótesis unilateral del tipo $H_0(\mu = \mu_0)$ vs $H_0(\mu > \mu_0)$, la expresión de la curva característica sería:

$$OC(d) = \Phi(Z_{\frac{\alpha}{2}} - d\sqrt{n}), \quad d \geq 0$$

y el tamaño de la muestra en este supuesto unilateral de la hipótesis alternativa se podría calcular fijando los puntos de la curva característica $(0, \alpha)$ y (d, β) mediante:

$$n \geq \left(\frac{Z_{\alpha} + Z_{\beta}}{d} \right)^2$$

ya que

$$\Phi(Z_{\frac{\alpha}{2}} - d\sqrt{n}) \leq \beta \Rightarrow -Z_{\beta} \geq Z_{\frac{\alpha}{2}} - d\sqrt{n}$$

Ejemplo: En la fabricación de botes de refrescos se quiere verificar que su valor nominal está entorno a los 33 cl que marca el envase, para determinar el valor de la curva característica para un valor del tamaño muestral de 5 se queremos detectar un cambio cuando el contenido medio sea inferior a 32 cl, suponiendo una desviación típica poblacional de $\sigma_0 = 1,23$ obtendríamos:

$$d = \frac{|\mu - \mu_0|}{\sigma_0} = \frac{|33 - 32|}{1,23} = 0,81301$$

Por tanto $OC(d) = \Phi(1,645 - 0,81301\sqrt{5}) = \Phi(-0,17) = 0,4325$:

Esto quiere decir que si la media del proceso fuese realmente 32 cl en el 43.25% de las muestras de tamaño 5 se llegaría a aceptar que la media real

es de 33 cl (Hipótesis nula), es decir cometeríamos un error de tipo II con probabilidad 0.4325, lo que llevaría a buscar otro test distinto o bien a aumentar el tamaño muestral con objeto de hacer este más potente (la potencia en este caso sería sólo del 56.75 % ($1 - 0,4325$)).

Si quisieramos una potencia del test del 95 % ($\beta = 0,05$, $Z_\beta = 1,645$)

$$n \geq \left(\frac{Z_\alpha + Z_\beta}{d} \right)^2 = \left(\frac{1,645 + 1,645}{0,81301} \right)^2 = 16,37$$

1.4. El ARL para gráficos de variables

El uso de las curvas características permite estimar el tamaño muestral necesario para detectar cambios en el riesgo β . La probabilidad de que se detecte un cambio en la k -ésima muestra posterior, será $1 - \beta$ veces la probabilidad de no descubrirlo en cada una de las $k - 1$ muestras iniciales. Así la probabilidad de detectar el cambio en la primera muestra es $1 - \beta$, en la segunda muestra será $\beta(1 - \beta)$, en general la probabilidad de conseguir detectar el cambio en la k -ésima muestra será: $\beta^{k-1}(1 - \beta)$. Teniendo esto en cuenta podemos estimar el número de submuestras que habría que analizar antes de detectar el cambio (se trataría de una distribución geométrica):

$$\sum_{k=1}^{\infty} k \beta^{k-1} (1 - \beta) = \frac{1}{1 - \beta}$$

A este valor del número de muestras necesarias para detectar la primera muestra defectuosa o fuera de control se denomina *ARL* (Average Run Length).

$$ARL = \frac{1}{1 - \beta}$$

El ARL o longitud media de racha es el promedio de puntos que hacen falta representar en el gráfico (muestras a tomar) hasta que el primer punto muestre señal de fuera de control. El valor del ARL puede calcularse analíticamente o por medio de simulación (véase la figura 1-5).

Si por ejemplo se tuviese un valor para β de 0.05 el número esperado de submuestras que habría que analizar para detectar el cambio sería de 1053, mientras que si el riesgo β fuese grande, por ejemplo 0.75, harían falta sólo 4 muestras. Es precisamente a raíz del estudio de este tipo de curvas OC cuando surgen las recomendaciones de tomar 20 o 25 submuestras para establecer los límites de control. En el caso en que se si verifica la hipótesis nula, es decir, cuando si está el proceso bajo control, se tiene:

$$ARL_{d=0} = \frac{1}{\alpha}$$

donde α es la probabilidad del error de primera especie. Si por ejemplo estamos utilizando gráficos de control 3 sigma, saldría un ARL de $1/0,0027 = 370,37$,

Para los gráfico de Shehart se pueden graficar distintos valores para el *ARL* en función de k y n .

El conocimiento de los valores de ARL para proceso bajo control y fuera de control es fundamental en el diseño de los gráficos de control de calidad. Pues debemos asegurarnos de tener una baja tasa de falsas alarmas (ARL bajo control alto) y, al mismo tiempo, rapidez suficiente para detectar los cambio (ARL fuera de control bajo).

También se puede utilizar el ATS (Average Time to Signal) o tiempo promedio hasta la señal, como el tiempo promedio que tardará el gráfico hasta que éste indique una señal de falta de control. Resulta sencillo obtener una fórmula aproximada para el valor de ATS si conocemos el tiempo entre muestras, h :

$$ATS = ARL.h$$

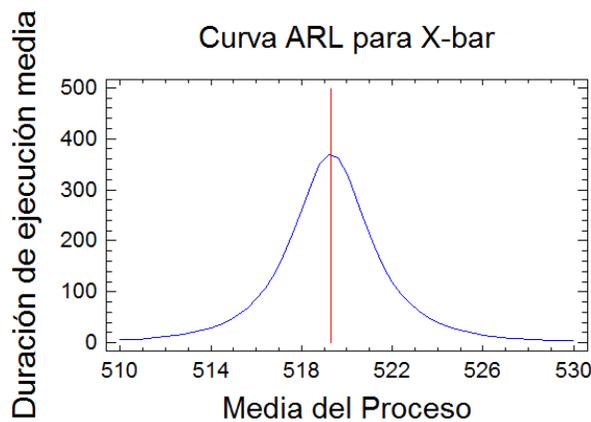


Figura 1-5 Curva ARL.

1.5. Ejercicios del tema

1.5.1. Ejercicio 1:

En un control de calidad la responsable del empaquetado está interesada en controlar el contenido de humedad de dichos paquetes, para ello seleccionada 32 submuestras de tamaño 5, como medida toma el contenido de humedad en los paquetes de sal, como porcentaje del peso total. Resultando una media de

las 32 submuestras de 2.6502, una media de las desviaciones típicas de 0.2425 y una media de los recorridos de 0.6066

- a) Calcular los límites de control utilizando las desviaciones al nivel 3σ .
- b) Calcular los límites de control utilizando los recorridos al nivel 3σ .
- c) A la vista de los resultados, ¿qué diagrama de los anteriores considera mejor?

1.5.2. Ejercicio 2

Para realizar un control de calidad sobre el porcentaje de materia grasa de una determinada marca de mantequilla, se seleccionaron 100 embases de mantequilla, elegidos en submuestras de tamaño 5. Se determinó el contenido de materia grasa total de cada embase, resultando los datos de la tabla adjunta.

- a) Calcula los límites de los gráficos de control \bar{X} de medias para el contenido medio de materia grasa.
- b) Calcula los límites de los gráficos de control de desviaciones típicas
- c) Considerando como intervalo de tolerancias de materia grasa [65,95], ¿cuál sería la capacidad del proceso según una metodología 3σ ?

Subgrupo Tamaño X-bar S

1	5	85,4626	8,21722
2	5	78,1281	3,66889
3	5	78,9408	3,2049
4	5	78,9221	5,43189
5	5	82,544	5,77852
6	5	79,9901	5,08758
7	5	78,8692	4,22067
8	5	78,8248	7,59128
9	5	79,5877	1,63957
10	5	78,5338	3,64375
11	5	78,8945	4,54403
12	5	80,7316	3,18696
13	5	76,208	3,83605
14	5	79,589	6,29734
15	5	83,5223	5,30842
16	5	80,1668	5,11653
17	5	77,278	3,28091
18	5	81,6929	1,22231
19	5	76,6554	5,13308
20	5	76,1949	3,79899

Pueden ser de interés los siguientes datos (calculados a partir de la tabla anterior):

Media del Proceso = 79,5368
Sigma del Proceso = 4,79847
Sigma Media = 4,51044

1.5.3. Ejercicio 3

Para controlar la calidad de un acero, se toman 10 muestras de tamaño 6 y se calculan la media y la cuasidesviación típica muestral de la dureza (en Rocwell C) de cada una de ellas. A continuación se adjuntan los resultados de las medias de cada muestra, indicando entre paréntesis la desviación típica: 53 (9.3), 51 (3.1), 48 (2.5), 59 (1.3), 48 (6.7), 51 (7.9), 49 (3.8), 51 (8.2), 48 (2.6), 57 (5.4).

- Calcular un intervalo de confianza, al 99.7%, para la dureza media del acero, puede utilizarse como desviación típica muestral 5.51.
- Deducir los límites de los gráficos de control \bar{X} de Medias y de desviaciones típicas. ¿Qué conclusiones se sacarían de cara a controlar la calidad? ¿Estaría el proceso bajo control?
- Considerando como amplitud del intervalo de tolerancia 30, ¿Cuál sería la capacidad del proceso según una metodología 3σ ?

1.5.4. Ejercicio 4

Se está investigando la consistencia de un proceso de fabricación. Los siguientes datos se tomaron en cuatro momentos distintos de un día. Cada vez, se extrajeron cinco artículos acabados y se obtuvieron las siguientes medidas para cada uno:

M 1	M 2	M 3	M 4
1.025	1.036	1.011	1.022
1.042	1.016	1.029	1.027
1.013	1.028	1.031	1.046
1.027	1.023	1.021	1.033
1.018	1.025	1.019	1.034

- Determinar los gráficos de control de la media (a partir de la desviación típica y del recorrido), la desviación típica y el recorrido para un nivel de confianza del 95%. Dibujar las cuatro muestras en estos gráficos.
- Discutir si se debe hacer algún cambio a los gráficos de control obtenidos.
- Estimar la proporción de artículos que tendrán una medida menor o igual que 1.000 cuando el proceso está bajo control.

Bibliografía

- [1] Aparisi, F. (1996). A comparison between the simultaneous use of Shewhart charts and the use of the T squared control chart. Proceedings of the Section on Physical and Engineering Sciences of the American Statistical Association. USA.
- [2] Bai, D. S. y Choi, I. S. (1995) X and R Control Charts for Skewed Populations. *Journal of Quality Technology*, 27, pp. 120-131.
- [3] Blazek, L. W., Novic B. y Scott M. D. (1987): Displaying Multivariate Data Using Polyplots. *Journal of Quality Techonlogy*. 19 (2) pp 69-74.
- [4] Chakraborti, S., Van Der Laan, P. y Bakir, S. T. (2001) Nonparametric Control Charts: An Overview and Some Results. *Journal of Quality Technology*, 33, pp. 304-315.
- [5] Doganaksoy, N., Faltin, F. W. y Tucker. W.T. (1991), Identification of out-of-control characteristics in a multivariate manufacturing environment, *Communications in Statistics: Theory and Methods*, Vol. 20, pp. 2775-2790.
- [6] Jones, L. A. y Woodall, W. H. (1998) The Perfomance of Bootstrap Control Charts. *Journal of Quality Technology*, 30, No 4, pp. 362-375.
- [7] Montgomery, D.C. (2001). *Introduction to Statistical Quality Control*. 4th edition John Wiley.
- [8] Shewhart, W.A. (1931). *Economic Control of Quality of Manufactured Product*, Republished in 1986 by Dover Publications.
- [9] Vermat, M. B., Ion, R., Does, R.J. y Klaassen, C. (2003) A Comparison of Shewhart Individuals Control Chart Based on Normal, Non-Normal, and Extreme Value Theory. *Quality and Reliability Engineering International*, 19, pp. 337-353.
- [10] Willeman, T. R. y Runger, G. C. (1996) Designing Control Charts Using an Empirical Reference Distribution. *Journal of Quality Technology*, 28, pp. 31-38.
- [11] Woodall, W.H. (1985). The statistical design of quality control charts. *The Statistician*, 34. pp. 155-160.