

# TAMAÑO Y FRECUENCIA DE MUESTREO EN GRÁFICOS DE CONTROL

CARMEN CAPILLA ROMÁ

RAFAEL ROMERO VILAFRANCA

Universidad Politécnica de Valencia

*Se analiza la elección óptima del tamaño y frecuencia de muestreo en gráficos de medias. Asumiendo que el coste total del muestreo es proporcional al número total de Uds muestreadas. El procedimiento óptimo depende de la magnitud de la desviación respecto del nominal que se considere relevante, que a su vez está relacionada con la capacidad del proceso. Se obtiene que, en general, a medida que aumenta dicha capacidad es aconsejable reducir el tamaño de la muestra e incrementar la frecuencia de muestreo.*

## Subgroup size and sampling interval in Control Charts

**Keywords:** Statistical Process Control, Process Capability, Shewhart Control Charts, Average Run Length

## 1. INTRODUCCIÓN

Un aspecto esencial en la moderna filosofía de la calidad es el énfasis sobre la prevención, frente a los enfoques tradicionales basados fundamentalmente en la detección a posteriori de los defectos mediante métodos de inspección. Este énfasis se ha traducido en una gran expansión de la utilización industrial de técnicas de control on-line, especialmente de los gráficos de control desarrollados por Shewart (1931) hace más de 50 años.

---

—Carmen Capilla Romá - Enginyer Agrònom. Dep. d'Estadística i I.O. de la Universitat Politècnica de València.

—Rafael Romero Villafranca - Catedràtic d' Universitat - Dept. d'Estadística - Universitat Politècnica de València.

—Article rebut el juny de 1989.

El objetivo de cualquier gráfico de control es detectar precozmente la presencia de causas especiales de variabilidad en un proceso. Implantados adecuadamente los gráficos de control pueden, adicionalmente, ayudar a identificar dichas causas especiales y eliminarlas del proceso manteniendo a éste bajo control.

Cuando la característica de calidad a controlar es una variable continua y la producción está constituida por gran número de unidades individuales o piezas, el tipo de gráfico más ampliamente difundido en la industria española es el  $\bar{X} - R$ . La utilización de este gráfico plantea como problemas previos la selección del tamaño de muestra a utilizar y la determinación de la frecuencia de muestreo.

En situaciones en las que el costo total del muestreo es aproximadamente proporcional al número total de unidades muestreadas, se plantea el dilema de elegir entre planes basados en muestras pequeñas tomadas a intervalos cortos frente a otros consistentes en muestras de mayor tamaño pero más espaciadas en el tiempo.

El presente trabajo analiza cómo dicha elección afecta a la potencia del gráfico definida ésta como su capacidad para detectar rápidamente las salidas de control del proceso. Dicho análisis permite, bajo ciertas hipótesis razonables, establecer el tamaño óptimo de muestra como función de la capacidad del proceso que se pretende controlar.

## 2. CONSIDERACIONES ESTADÍSTICAS PREVIAS

Desde el punto de vista estadístico un gráfico de control no es más que una herramienta para realizar periódicamente contrastes de hipótesis respecto a la distribución de probabilidad del resultado del proceso. En particular el gráfico  $\bar{X}$  se utiliza para detectar la presencia de causas especiales de variabilidad que modifican la media de dicha distribución.

Las propiedades estadísticas de un contraste de este tipo vienen caracterizadas por sus riesgos de primera y segunda especie. Como es bien sabido el riesgo de primera especie  $\alpha$  no es más que la probabilidad de obtener una falsa señal de falta de control, es decir una señal de falta de control que se produce pese a que la media del proceso no se ha modificado, mientras que el riesgo de segunda especie  $\beta$  está asociado a la probabilidad de no detectar las salidas de control del proceso.

Intimamente relacionados con los riesgos  $\alpha$  y  $\beta$  se manejan en la teoría del control de procesos los conceptos de ARL (Average Run Length) y AvD (Average Duration).

El ARL es el valor medio del número de muestras que hay que tomar antes

de que aparezca una señal de falta de control. Se deduce muy fácilmente que el ARL no es más que el inverso de la probabilidad de aparición de una señal de falta de control, resultando por tanto igual a  $1/\alpha$  para un proceso bajo control.

Cuando se comparan planes de control con diferentes frecuencias de muestreo resulta relevante el concepto de AvD, que no es más que el valor medio del tiempo que se tarda en detectar una salida de control del proceso. Denominando  $f$  al intervalo de tiempo entre dos muestras consecutivas, el AvD está relacionado con el ARL mediante la expresión (Ewan et al., 1960).

$$\text{AvD} = (\text{ARL} - 0.5) \times f$$

que asume que la salida de control se produce, en promedio, en el punto central del intervalo entre dos muestras consecutivas.

Con carácter general, para un riesgo de primera especie  $\alpha$  dado, la probabilidad de detectar una salida de control aumenta, y disminuye en consecuencia el ARL, a medida que aumenta el tamaño  $n$  de la muestra. Adicionalmente el AvD disminuirá, obviamente, a medida que lo haga el intervalo de muestreo  $f$ . Por otra parte, tanto el ARL como el AvD disminuyen cuanto más marcada es la salida de control del proceso, es decir cuanto mayor resulta, medida en unidades de desviación típica  $\sigma$  del proceso, la diferencia entre la nueva media y la existente bajo control.

La elección de un determinado plan de muestreo en un gráfico de control debe basarse en consideraciones tanto económicas como estadísticas. Las primeras permiten determinar el esfuerzo total de muestreo a realizar en función de parámetros como el coste unitario de muestreo, la probabilidad de salidas de control y el coste de las unidades defectuosas producidas antes de que éstas se detecten. Las segundas ayudan a distribuir dicho esfuerzo de forma que se optimicen las propiedades estadísticas del gráfico.

En particular, asumiendo que el coste total del muestreo es proporcional al número de unidades muestreadas es posible estudiar cuáles deben ser los valores de  $n$  y  $f$  que minimizan el AvD en función de la magnitud de la salida de control del proceso, según se analiza en los siguientes apartados.

### **3. RELACIÓN DEL AvD CON EL TAMAÑO Y FRECUENCIA DE MUESTREO**

Los procedimientos de muestreo que se van a comparar vienen definidos por un tamaño de muestra  $n$  y un intervalo de muestreo  $f$  tales que el número total de artículos muestreados por unidad de tiempo es el mismo en todos los casos. Admitiendo la simplificación de suponer el coste de muestreo propor-

cional a este último número, todos los procedimientos resultan en consecuencia equivalentes desde el punto de vista de costo.

El intervalo de muestreo  $f$  viene expresado en todos los casos como múltiplo de una unidad de tiempo arbitraria.

Con el fin de cubrir la gama de situaciones más frecuentes en la práctica industrial habitual se van a comparar los 4 procedimientos siguientes:

- Procedimiento 1 : Tamaño  $n=1$  cada  $f = 1$  Ud. de tiempo  
 "            2 : Tamaño  $n=3$  cada  $f = 3$             "  
 "            3 : Tamaño  $n=5$  cada  $f = 5$             "  
 "            4 : Tamaño  $n=7$  cada  $f = 7$             "

Para el cálculo del AvD se ha utilizado la fórmula ya indicada:  $AvD = (ARL - 0.5) \times f$ , donde ARL es igual a  $1/p$ , siendo  $p$  la probabilidad de que se produzca una señal de falta de control.

En el presente trabajo se ha considerado como única señal de este tipo la aparición de un punto fuera de los límites de control del gráfico  $\bar{X}$ . También se asume, de la forma habitual, que la característica a controlar se distribuye normalmente.

Con el fin de obtener procedimientos comparables se ha ajustado la posición de los límites de control de forma que el tiempo promedio (AvD) que tarda en producirse una falsa alarma, es decir una señal de falta de control no habiéndose modificado la media del proceso, sea el mismo en todos los casos. Concretamente todos los AvD bajo control se han hecho iguales a los de un gráfico con tamaño  $n = 5$  cada  $f = 5$  unidades de tiempo, y límites de control estándar  $m \pm 3\sigma/\sqrt{5}$ .

Para este procedimiento, cuando el proceso está bajo control, se tiene:

$$ARL = 1/\alpha = 370.37$$

siendo  $\alpha$  el riesgo de 1ª especie 2.7 ‰, con lo que aplicando la fórmula correspondiente al AvD:

$$AvD = (370.37 - 0.5) \times 5 \approx 1850 \text{ Ud. de tiempo}$$

Para cualquier otro procedimiento se debe tener por tanto bajo control:

$$AvD = 1850 = (ARL - 0.5) \times f$$

de donde puede despejarse el valor del ARL:

$$ARL = (1850/f) + 0.5 = 1/p$$

siendo  $p$  la probabilidad de obtener un punto fuera de los límites de control cuando el proceso se halla bajo control. Este resultado permite calcular fácilmente la posición de los límites de control para los restantes procedimientos, obteniéndose los resultados que se reflejan a continuación.

$$\begin{aligned} n = 1 \quad f = 1 \quad m \pm 3.46 \sigma & \quad (\alpha = 0.54\% \omega) \\ n = 3 \quad f = 3 \quad m \pm 3.15 \sigma / \sqrt{3} & \quad (\alpha = 1.62\% \omega) \\ n = 5 \quad f = 5 \quad m \pm 3 \sigma / \sqrt{5} & \quad (\alpha = 2.7 \% \omega) \\ n = 7 \quad f = 7 \quad m \pm 2.9 \sigma / \sqrt{7} & \quad (\alpha = 3.77\% \omega) \end{aligned}$$

Si la media del proceso sufre una desviación igual a  $\delta\sigma$  respecto del valor nominal  $m$  el proceso estará operando a un nuevo nivel  $m + \delta\sigma$ . El ARL correspondiente a esta situación, para un procedimiento consistente en tomar muestras de tamaño  $n$  cada  $f$  unidades de tiempo es igual a  $1/p$ , siendo  $p$  en este caso la probabilidad de que la media muestral caiga fuera de los límites de control  $m \pm h\sigma/\sqrt{n}$ . Dicha probabilidad viene dada por la expresión:

$$\begin{aligned} 1/\text{ARL} = & P(N(m + \delta\sigma, \sigma^2/n) > m + h\sigma/\sqrt{n}) + \\ & + P(N(m + \delta\sigma, \sigma^2/n) < m - h\sigma/\sqrt{n}) \end{aligned}$$

y conociendo el ARL, el tiempo promedio que tarda en detectarse la salida de control se puede calcular a partir de la fórmula dada anteriormente para el AvD.

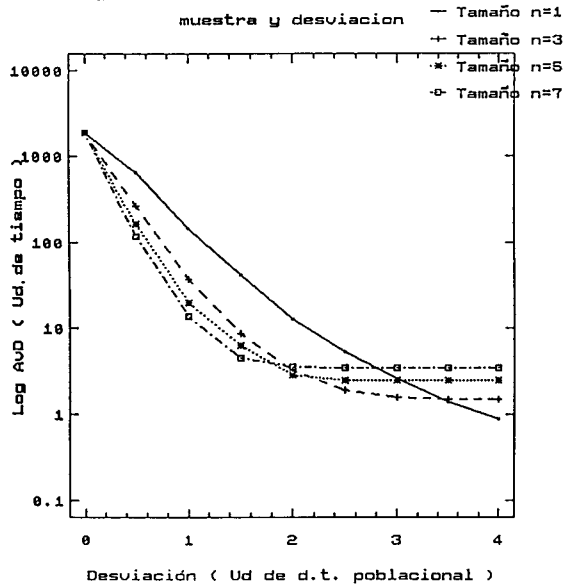
Los resultados obtenidos para el valor del AvD en función del procedimiento de muestreo utilizado y de la magnitud  $\delta$  de la salida de control se reflejan en la Tabla 1 y en la correspondiente Figura 1 (En esta última se utiliza escala logarítmica para los valores del AvD).

Como puede apreciarse el plan de muestreo óptimo, entendiéndose como tal el que minimiza el AvD, depende de la magnitud de la modificación de la media del proceso.

Tabla 1 : AvD (Ud. de tiempo)

Desviación (Ud. de $\sigma$ )	n=1 f=1	n=3 f=3	n=5 f=5	n=7 f=7
0.0	1850.0	1850.0	1850.0	1850.0
0.5	649.0	263.0	164.0	117.0
1	145.0	37.1	19.9	13.9
1.5	42.4	8.8	6.4	4.6
2	13.4	3.3	2.9	3.6
2.5	5.4	1.9	2.5	3.5
3	2.6	1.6	2.5	3.5
3.5	1.4	1.5	2.5	3.5
4	0.9	1.5	2.5	3.5

Figura 1 - AvD en función del tamaño de muestra y desviación



En el gráfico y tabla anteriores se observa, en efecto, que cuando la desviación respecto del valor nominal es inferior a  $1.5 \sigma$  el plan más efectivo es el basado en muestras de tamaño 7 tomadas cada 7 unidades de tiempo, mientras que, en el otro extremo, para detectar precozmente desviaciones superiores a  $3.5 \sigma$  la mejor solución consiste en tomar muestras de tamaño 1 cada unidad de tiempo.

El plan a adoptar depende por tanto del orden de magnitud, expresado en unidades de desviación típica del proceso, de las desviaciones respecto al nominal cuya detección precoz resulte relevante. En definitiva el plan óptimo está estrechamente relacionado con la capacidad del proceso a controlar, según se analiza en el siguiente apartado.

#### 4. RELACIÓN CON LA CAPACIDAD DEL PROCESO

Como es bien sabido el índice de capacidad potencial de un proceso  $C_p$  mide la relación existente entre la anchura de especificación y la variabilidad natural del proceso, definida esta última como seis veces la desviación típica del mismo

$$C_p = (LES - LEI)/6\sigma$$

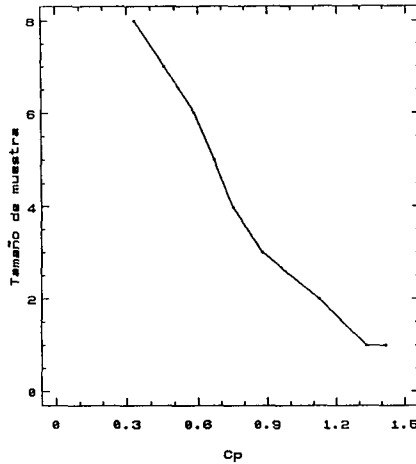
siendo LES y LEI los límites de especificación superior e inferior para las características en cuestión.

En un proceso centrado el valor del  $C_p$  coincide con el del índice de capacidad real del proceso  $C_{pk}$  y permite obtener el porcentaje de piezas fuera de especificación producidas cuando el proceso se halla bajo control. En esta situación la distancia de la media del proceso a los límites de especificación, medida en unidades de desviación típica, no es más que el triple del valor del  $C_p$ .

Un objetivo razonable para establecer un plan de muestreo es intentar detectar rápidamente salidas de control que conduzcan a un incremento notable en la proporción de unidades producidas fuera de especificación. Concretamente vamos a asumir que se trata de minimizar el AvD cuando la desviación respecto del valor nominal es tal que produce un porcentaje de defectuosos superior o igual al 50%.

La hipótesis anterior implica que el plan de muestreo óptimo es el que minimiza el AvD para una desviación  $\delta\sigma$  en la que  $\delta$  no es más que  $3 \times C_p$ . En la Figura 2 se ha reflejado el tamaño de muestra óptimo, asumiendo como siempre constancia en el número total de unidades muestreadas, en función de la capacidad  $C_p$  del proceso.

FIGURA 2. Tamaño de muestra óptimo.



En la Figura se aprecia que cuando el proceso opera con un índice de capacidad  $Cp$  alto ( $> 1.33$ ) el procedimiento de control óptimo consiste en tomar muestras de tamaño  $n = 1$  cada unidad de tiempo, mientras que, en el otro extremo, para capacidades de proceso  $Cp$  bajas el procedimiento óptimo consiste en tomar muestras de tamaño grande disminuyendo la frecuencia del muestreo.

El valor  $n = 5$ , posiblemente el más frecuente en la industria española, resulta adecuado para procesos con índices de capacidad comprendidos entre 0.6 y 0.8.

El resultado obtenido para procesos con capacidad superior a 1.33 coincide en cierto sentido con la propuesta de control óptimo de Taguchi (1981) desarrollada en su obra "On-line Quality Control during Production", lo que no es sorprendente dada la en general elevada capacidad de los procesos en la industria japonesa.

## 5. CONSIDERACIONES FINALES

Los resultados obtenidos en el presente trabajo indican que, bajo los supuestos considerados en el mismo, el control mediante gráficos  $\bar{X} - R$  de procesos de capacidad satisfactoria debería realizarse utilizando tamaños de muestra sensiblemente inferiores a los usados habitualmente en la industria. Con capacidades superiores a 1.3, el recurso a muestras individuales tomadas con mayor frecuencia resulta la solución óptima desde el punto de vista de minimizar el AvD para salidas de control importantes, considerando como tales a las que llevan a producir más de un 50% de piezas fuera de especificación. Por otro



lado, un inconveniente de la utilización de muestras individuales es la imposibilidad de detectar cambios en la dispersión del proceso, a menos que se lleve en paralelo un gráfico de recorridos móviles. El estudio en este caso del tamaño de agrupamiento óptimo para distintas desviaciones respecto de la media, se desarrolla en otro trabajo de los autores.

Los resultados expuestos sólo son aplicables en aquellas situaciones en las que el costo total de muestreo es aproximadamente proporcional al número total de unidades muestreadas.

Fijando el tamaño de muestra la determinación del intervalo de muestreo óptimo, es decir la fijación de la unidad de tiempo que a lo largo del presente trabajo se ha considerado como arbitraria, debe basarse fundamentalmente en consideraciones de tipo económico, especialmente en los costos de muestreo y de las unidades defectuosas producidas y en la probabilidad y magnitud de las posibles salidas de control del proceso. Una interesante aportación en este sentido puede encontrarse en la referencia mencionada de Taguchi.

## 6. AGRADECIMIENTO

Los autores desean manifestar su agradecimiento al Banco de Vizcaya, actualmente Banco Bilbao-Vizcaya, por la beca concedida para la realización, entre otros, del presente trabajo.

## 7. BIBLIOGRAFÍA

- [1] **Duncan, A.K.** (1986). "Quality Control and Industrial Statistics". Quinta Edición, Richard D. Irwin, Homewood, Illinois.
- [2] **Ewan, W.D. and Kemp, K.W.** (1960). "Sampling Inspection of Continuous Processes with no Autocorrelation Between Successive Results". *Biométrica* 47, pp 363-380.
- [3] **Shewart, W.A.** (1931). "Economic Control of Quality of Manufactured Product". Van Nostrand, New York.
- [4] **Taguchi, G.** (1981). "On line Quality Control During Production". Japanese Standards Association, Tokyo, Japan.

