

**UNIVERSIDAD POLITÉCNICA DE VALENCIA**  
**DEPARTAMENTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA**

***APORTACIONES AL CONOCIMIENTO DE  
LAS PROPIEDADES ESTADÍSTICAS DE LOS  
GRÁFICOS DE CONTROL PARA LA MEDIA***

**CARMEN CAPILLA ROMÁ**  
**Valencia, Diciembre 1991**

A mis padres

## **AGRADECIMIENTOS**

Quiero expresar mi agradecimiento a las instituciones y personas que han hecho posible el presente trabajo de investigación. No pretendo realizar una enumeración exhaustiva de todas ellas, pues a buen seguro que alguna se quedará en el tintero.

En primer lugar a Rafael Romero Villafranca por la labor de dirección de la misma. Además por la confianza puesta en mi para la realización de esta investigación.

A todos los profesores del Departamento de Estadística e Investigación Operativa por lo grata que han hecho mi tarea y por la formación proporcionada. En especial quiero mencionar a Concepción Maroto Álvarez, por su ejemplo de compañerismo y por la ayuda prestada en la fase de redacción de la presente tesis.

Por último a la Consellería de Cultura, Educación y Ciencia y a la Fundación del Banco Bilbao-Vizcaya por la financiación de este trabajo.

## ÍNDICE

<b>CAPÍTULO 1- ANTECEDENTES Y OBJETIVOS</b> .....	8
<b>1.1- INTRODUCCIÓN</b> .....	8
<b>1.2- ANTECEDENTES BIBLIOGRÁFICOS</b> .....	16
1.2.1-Diseño óptimo de gráficos de medias.....	16
1.2.2-Señales adicionales de falta de control en gráficos de me- dias.....	23
1.2.3-Gráficos CUSUM.....	30
1.2.4-Gráficos de medias móviles.....	41
1.2.5-Gráficos de control y datos correlacionados.....	43
<b>1.3- OBJETIVOS</b> .....	48
<b>CAPÍTULO 2- METODOLOGÍA</b> .....	54
<b>2.1- COMPARACIÓN DE GRÁFICOS DE CONTROL: ARL</b> .....	54
<b>2.2- PROCEDIMIENTOS PARA CÁLCULO DEL ARL</b> .....	59
<b>2.3- GENERACIÓN DE OBSERVACIONES</b> .....	61
<b>2.4- ANÁLISIS DEL VECTOR BÁSICO DE PERTURBACIONES</b> .....	62
<b>2.5- DETERMINACIÓN DEL NÚMERO DE REPETICIONES</b> ..	64
<b>2.6- SIMULACIÓN DEL FUNCIONAMIENTO DE LOS DISTINTOS GRÁFICOS</b> .....	68
2.6.1-Simulación de la salida de control.....	68
2.6.2-Simulación del gráfico de medias.....	69
2.6.3-Simulación del gráfico de medias móviles.....	70
2.6.4-Simulación del gráfico CUSUM.....	70
2.6.5-Simulación del gráfico EWMA.....	73
2.6.6-Simulación de observaciones correlacionadas .....	76

<b>CAPÍTULO 3- FRECUENCIA Y TAMAÑO DE MUESTREO EN GRÁFICOS DE MEDIAS.....</b>	<b>77</b>
<b>3.1- INTRODUCCIÓN.....</b>	<b>77</b>
<b>3.2- CONSIDERACIONES ESTADÍSTICAS PREVIAS.....</b>	<b>78</b>
<b>3.3- RELACIÓN DEL A<sub>v</sub>D CON EL TAMAÑO Y FRECUENCIA DE MUESTREO.....</b>	<b>83</b>
<b>3.4- RELACIÓN ENTRE EL PLAN DE MUESTREO Y LA CAPACIDAD DEL PROCESO.....</b>	<b>88</b>
<b>CAPÍTULO 4- SEÑALES DE FALTA DE CONTROL EN GRÁFICOS DE MEDIAS .....</b>	<b>94</b>
<b>4.1- INTRODUCCIÓN.....</b>	<b>94</b>
<b>4.2- CONSIDERACIONES PREVIAS.....</b>	<b>96</b>
<b>4.3- DISEÑO DE GRÁFICOS DE CONTROL A COMPARAR.....</b>	<b>98</b>
<b>4.4- EFECTO DE LAS SEÑALES ADICIONALES SOBRE EL ARL.....</b>	<b>103</b>
4.4.1-Desviación constante en la media.....	104
4.4.2-Desviación progresiva en la media.....	108
<b>4.5- FUNCIÓN DE PESO DE LOS PROCEDIMIENTOS COMPARADOS.....</b>	<b>110</b>
<b>4.6- FRECUENCIA RELATIVA DE LAS SEÑALES DE FALTA DE CONTROL.....</b>	<b>112</b>
4.6.1-Desviación constante en la media.....	113
4.6.2-Deriva progresiva en la media.....	117
4.6.3-Frecuencia de las señales de falta de control cuando aumenta la dispersión del proceso.....	120
<b>4.7-OTRAS SEÑALES DE FALTA DE CONTROL.....</b>	<b>122</b>
<b>CAPÍTULO 5- GRÁFICOS DE CONTROL PARA OBSERVACIONES INDIVIDUALES.....</b>	<b>127</b>

<b>5.1- INTRODUCCIÓN.....</b>	<b>127</b>
<b>5.2- PROPIEDADES ESTADÍSTICAS DE LOS GRÁFICOS DE MEDIAS MÓVILES.....</b>	<b>129</b>
5.2.1-Consideraciones previas.....	129
5.2.2-Efecto del tamaño de agrupamiento en el ARL.....	134
5.2.3-Rachas en el gráfico de medias móviles.....	140
<b>5.3- COMPARACIÓN DE GRÁFICOS DE CONTROL PARA OBSERVACIONES INDIVIDUALES.....</b>	<b>145</b>
 <b>CAPÍTULO 6- GRÁFICOS DE CONTROL Y DATOS CORRELACIONA- DOS.....</b>	 <b>151</b>
<b>6.1- INTRODUCCIÓN.....</b>	<b>151</b>
<b>6.2- CONSIDERACIONES PREVIAS.....</b>	<b>153</b>
<b>6.3- EFECTO DE LA CORRELACIÓN EN EL ARL.....</b>	<b>157</b>
6.3.2- Gráfico de medidas individuales.....	157
6.3.3- Gráfico de medias móviles.....	159
6.3.4- Gráfico CUSUM.....	166
6.3.5- Gráfico EWMA.....	170
<b>6.4- CONSIDERACIONES ADICIONALES.....</b>	<b>174</b>
 <b>CAPÍTULO 7- CONCLUSIONES.....</b>	 <b>177</b>
<b>7.1- TAMAÑO Y FRECUENCIA DE MUESTREO EN GRÁFI- COS DE MEDIAS.....</b>	<b>177</b>
<b>7.2- SEÑALES DE FALTA DE CONTROL EN GRÁFICOS DE MEDIAS.....</b>	<b>177</b>
<b>7.3- GRÁFICOS DE CONTROL PARA MEDIDAS INDIVIDUA- LES.....</b>	<b>178</b>

7.4- GRÁFICOS DE CONTROL Y DATOS CORRELACIONA- DOS.....	180
CAPÍTULO 8- REFERENCIAS BIBLIOGRÁFICAS.....	182
ANEJO 1- GRÁFICOS DE CONTROL PARA LA MEDIA.....	196
A.1.1- GRÁFICOS DE MEDIAS DE SHEWHART.....	196
A.1.2- GRÁFICO DE MEDIDAS INDIVIDUALES.....	200
A.1.3- GRÁFICO DE MEDIAS MÓVILES.....	201
A.1.4- GRÁFICO CUSUM.....	202
A.1.5- GRÁFICO EWMA.....	206
ANEJO 2- PROGRAMAS UTILIZADOS EN LAS SIMULACIONES.....	210

# **CAPÍTULO 1**

## **ANTECEDENTES Y OBJETIVOS**

### **1.1- INTRODUCCIÓN**

La década de los ochenta ha visto nacer en la industria occidental un interés creciente respecto a los temas de control y mejora de la calidad. Este interés está íntimamente ligado a la comprensión de la relación existente entre la calidad y productividad, así como del papel fundamental que los esfuerzos en estos campos tienen sobre el incremento de la competitividad de las empresas en cualquier sector.

El impulso experimentado por estos temas en Europa y EEUU es debido, en gran parte, a los resultados espectaculares obtenidos en Japón, país que ha llevado a sus últimas consecuencias una filosofía cuyo principio básico es la mejora continua de la calidad y de la productividad en todos los procesos. La experiencia de las empresas de éxito muestra que calidad y productividad van ligadas en la práctica y que es posible aumentar ambas, reduciendo al mismo tiempo los costes de fabricación, y que la estrategia a aplicar no requiere grandes inversiones, sino la implantación de una filosofía operativa que se basa fundamentalmente en el aprovechamiento al máximo del capital humano y de la información potencial generada por todo proceso.

El trasplante y la adaptación de la filosofía y de las técnicas japonesas han llegado a la industria occidental de la mano fundamentalmente de las multinacionales de sectores, como el automovilístico o el electrónico, fuertemente afectadas por la competencia de aquel país. Estas grandes empresas han impulsado, por otra parte, la adopción de estas técnicas en sus proveedores, con lo que actualmente se asiste a un gran desarrollo en la aplicación de las mismas.



Una de las enseñanzas de las empresas japonesas es que una política racional sobre calidad no debe centrar sus esfuerzos en la inspección de los productos fabricados. La inspección no contribuye a la mejora de la calidad de los productos y procesos, sino que se limita a certificar la que ya existe. Por tanto, 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.

Juran et al. (1990) proporcionan algunos ejemplos de la evolución de los costes debidos a la calidad en función del tiempo, según se pasa de una situación en la que no se actúa en este sentido hasta un estado en el que se implanta un enfoque preventivo, pasando por la situación tradicional de énfasis en la inspección.

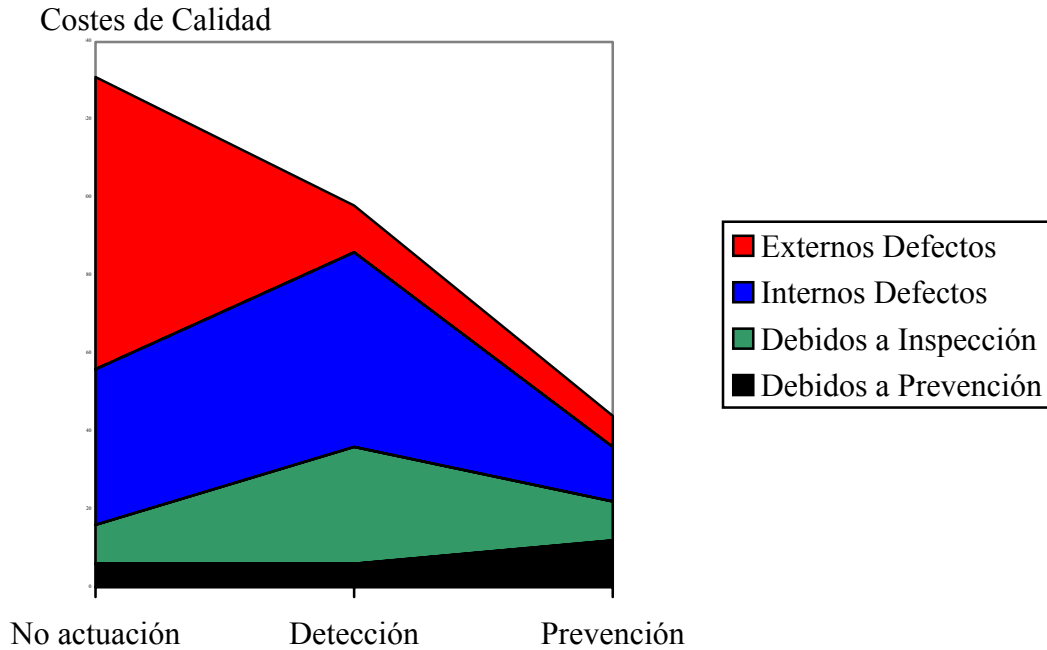
Se recoge en la FIGURA 1.1 una representación genérica de esta evolución en función de la estrategia desarrollada por la empresa. Según se observa, el esfuerzo preventivo consigue reducir los costes por producto defectuoso tanto internos (reprocesado, tiempo de paro, etc...), y los costes debidos a la verificación (inspección, ensayo, control del material de medición, etc....)

Un ingrediente fundamental en el enfoque preventivo del control de calidad es la utilización del método científico para la extracción eficiente de la información potencial generada por todo proceso, con el fin último de una toma racional de decisiones que permita minimizar la variabilidad presente en el mismo. Esto implica, entre otras cosas, el recurso a un conjunto de técnicas estadísticas, que proporcionan la metodología adecuada para la recogida y análisis de datos y que permiten evaluar las actuaciones sobre el proceso y el estado del mismo en cualquier momento.

En este sentido, el análisis on-line de los procesos bajo una óptica de prevención constituye la herramienta fundamental para compatibilizar niveles elevados de calidad y de productividad. Las técnicas estadísticas, aplicables con este fin, son la base de lo que se denomina **Control Estadístico de Procesos (SPC)**. Por tanto, es evidente el interés

que tienen las investigaciones sobre las propiedades de estos métodos, para determinar su utilización óptima en los diferentes contextos en que han de ser implantados.

FIGURA 1.1- Evolución de los costes debidos a la calidad.



Por otro lado, la mayoría de los problemas de calidad que se presentan en los procesos pueden resolverse aplicando métodos estadísticos muy sencillos. Entre ellos cabe destacar los **gráficos de control**, introducidos por Shewhart (1931) y que constituyen la herramienta fundamental de apoyo a los enfoques de prevención. Implantados adecuadamente, estos gráficos permiten detectar precozmente la presencia de causas especiales de variabilidad en un proceso y ayudar a su identificación y posterior eliminación, manteniendo así el proceso bajo control.

Cuando se da el caso, muy frecuente en industrias como la mecánica, electrónica, etc., de que la característica de calidad se puede medir en una escala continua y la producción está constituida por un gran número de unidades individuales, el procedimiento estándar más ampliamente difundido para controlar la posición del proceso es el **gráfico de medias muestrales de Shewhart**.

Otra solución estándar es el **gráfico de medidas individuales**, aplicable por ejemplo en el caso de la industria química o para control de procesos en los que se dispone de una sola medida por unidad de tiempo, y por lo tanto no tiene sentido hablar de tamaño de muestra. Como alternativa a este último, se propone con frecuencia la utilización del **gráfico de medias móviles**, en el que adoptando un determinado tamaño de agrupamiento, los valores observados individualmente se agrupan para calcular medias, que son a su vez registradas y analizadas según el procedimiento de medias muestrales estándar.

Existen además de los gráficos anteriormente mencionados, otros procedimientos alternativos, no tan extendidos en la práctica como los primeros, pero que pueden resultar más eficientes en ciertos contextos. Dentro de este grupo cabe distinguir los gráficos de Sumas Acumuladas, también llamados **gráficos CUSUM** y los **gráficos EWMA**, o gráficos de medias móviles exponencialmente ponderadas. Estos últimos son frecuentemente recomendados cuando se trata de controlar procesos en la industria química.

En el Anejo 1 se puede encontrar una explicación detallada sobre la construcción y fundamento de todos estos procedimientos.

Una encuesta realizada recientemente sobre la utilización de los métodos estadísticos por empresas pertenecientes a la Asociación Española para la Calidad (Peña et al., 1990) pone de manifiesto que estas empresas se encuentran en general en una fase inicial en el uso de estas técnicas. En concreto, en su mayoría el énfasis está puesto en la fase de detección; el control estadístico de procesos se encuentra en general poco desarrollado, y dentro de este conjunto de herramientas son los **gráficos de Shewhart** los más implantados, resultando prácticamente desconocidos los **gráficos CUSUM** y **EWMA**.

Nuestra investigación se ha desarrollado en un contexto de estrecho contacto con industrias de distintos sectores (automovilístico, químico, ...), con las que el Departamento mantiene **convenios de colaboración** a nivel de formación y

asesoramiento en la utilización de estas técnicas. En muchos casos dichos métodos se encuentran en la fase inicial de implantación, surgiendo a menudo cuestiones relacionadas con sus propiedades estadísticas, que deben ser analizadas con el fin de conseguir una utilización eficiente.

Por otro lado, la colaboración con empresas del sector químico, que difieren fundamentalmente de las del sector automovilístico o electrónico en cuanto al tipo de productos y procesos a controlar, ha llevado a cuestionarse hasta qué punto los procedimientos recomendados o estándar resultan eficientes en estos contextos.

En los procesos químicos, algunas de las hipótesis que asumen de partida los procedimientos estándar, como la independencia de observaciones, pueden dejar de cumplirse. En efecto, por tratarse en general de procesos continuos, simplemente el sistema de control automático implantado puede provocar que medidas tomadas en instantes de tiempo no muy alejados estén correlacionadas.

Resulta cuanto menos sorprendente la carencia de estudios sobre algunas cuestiones relacionadas con la utilización eficiente de estos procedimientos. Nuestra investigación trata de cubrir esa laguna. Creemos que su interés estriba en que se van a tratar aspectos relacionados con la utilización práctica de los gráficos de control más frecuentes, con el fin de estudiar sus propiedades y comportamiento, y que no han sido abordados hasta la fecha.

Teniendo en cuenta todo lo anterior, los **objetivos** de la presente investigación son los siguientes:

- 1) Analizar la elección del tamaño y frecuencia de muestreo en gráficos de medias muestrales, en función de la magnitud de desviación y de la capacidad del proceso.
- 2) Estudiar las propiedades estadísticas del gráfico de medias cuando se introducen señales adicionales de falta de control, comparando con el

procedimiento resultante con el gráfico CUSUM y el procedimiento estándar. Se considerarán los casos de desviación constante y progresiva en la media del proceso.

- 3) Determinar las propiedades estadísticas del gráfico de medias móviles en función del tamaño de agrupamiento. Comparar su efectividad con la correspondiente al gráfico de medidas individuales, CUSUM y EWMA.
- 4) Analizar el comportamiento estadístico de los gráficos de medidas individuales, medias móviles, CUSUM y EWMA cuando las observaciones individuales están correlacionadas.

La imposibilidad de determinar de forma analítica las propiedades de los gráficos en alguna de las cuestiones anteriores ha hecho necesario el recurso a técnicas de simulación en ordenador.

La presente tesis se ha estructurado en ocho capítulos. Tras esta introducción, en el **capítulo 1** se ofrece una visión panorámica de las referencias relacionadas con cada uno de los temas anteriores. Se revisan críticamente las principales líneas de investigación y la metodología empleada. A continuación, se sintetizan los resultados obtenidos por otros autores en el diseño óptimo de planes de muestreo, señales adicionales de falta de control, propiedades estadísticas de gráficos de medias móviles, determinación de propiedades estadísticas del gráfico CUSUM y el comportamiento de los gráficos de control cuando las observaciones están correlacionadas.

El **capítulo 2** está dedicado a fundamentar y describir la metodología utilizada, especificando todos los pasos seguidos en el procedimiento de simulación. Así, en primer lugar se discute la utilización del ARL como parámetro para comparación de los gráficos de control y los procedimientos a aplicar para su cálculo en cada caso. Se explica el método seguido para generar las observaciones y la elección del tamaño de simulación adoptado, y se valida comparando los resultados de algunas pruebas con los

correspondientes a la aplicación de métodos analíticos también aproximados. Finalmente, se expone el esquema general seguido para la simulación de cada gráfico.

En el **capítulo 3** se presentan los resultados, obtenidos en este caso analíticamente, relativos al tamaño y frecuencia óptimos cuando se produce una desviación de magnitud constante en la media, y en aquellas situaciones en las que el coste de muestreo es aproximadamente proporcional al número de artículos muestreados. Adicionalmente, se analiza la elección de este plan en función de la capacidad del proceso.

Los resultados del estudio de las propiedades estadísticas del gráfico de medias con señales adicionales se recogen en el **capítulo 4**. Se analiza su comportamiento en función del ARL, y se compara con el gráfico de medias estándar y el CUSUM, ajustados previamente con el fin de tener en los tres procedimientos un mismo número de falsas alarmas bajo control. Las salidas de control simuladas consisten en desviaciones constantes y progresivas en la media del proceso. Finalmente, se cuantifica la frecuencia relativa de aparición de cada señal de falta de control con el fin de precisar la conveniencia de su utilización.

En el **capítulo 5** se presentan los resultados del estudio de las propiedades estadísticas del gráfico de medias móviles. Se analiza su efectividad frente a distintas desviaciones en la media en función del tamaño de agrupamiento elegido y se cuantifica la cantidad de rachas en cada caso. A continuación se compara con otros procedimientos habituales en el contexto de observaciones individuales, como son los gráficos de medidas individuales, CUSUM y EWMA.

El **capítulo 6** recoge los resultados del estudio del comportamiento estadístico de los gráficos analizados en el capítulo anterior, pero cuando las observaciones individuales presentan correlación. Se considera el caso frecuente en la práctica de un proceso autorregresivo de primer orden AR(1) para las medidas de la característica de calidad.

En el **capítulo 7** se sintetizan los principales resultados obtenidos en esta tesis y se realizan una serie de consideraciones finales en torno a las conclusiones del estudio.

En la sección de bibliografía se recogen todas las referencias utilizadas en la elaboración de la presente tesis. Se incluyen algunas a las que no hemos podido acceder más que a través de recensiones de los trabajos originales, y otras de interés general sobre los temas tratados.

Los anejos recogen una síntesis sobre la construcción de los distintos gráficos de control analizados y los programas utilizados en su simulación.

## **1.2- ANTECEDENTES BIBLIOGRÁFICOS**

### **1.2.1- Diseño óptimo de gráficos de medias**

Numerosos autores han tratado el fundamento y utilización de las técnicas de control on-line de procesos. Exposiciones sobre los procedimientos más utilizados pueden encontrarse en Ott (1975), Ishikawa (1976), Johnson et al. (1977), Grant et al. (1980), Duncan (1986), Wadsworth et al. (1986), Ryan (1989), Juran et al. (1990), entre otros muchos.

Dentro de este conjunto de técnicas, los gráficos de control, introducidos en su versión estándar por Shewhart en 1931, se han convertido en una de las herramientas más importantes para control y mejora de procesos. El objetivo fundamental de estos gráficos es la detección on-line de las causas especiales de variabilidad en el proceso, con el fin último de identificarlas y ayudar a su posterior eliminación.

De los gráficos de control propuestos inicialmente por Shewhart, cuando la característica de calidad puede medirse en una escala continua, los procedimientos más utilizados son los gráficos de medias muestrales y recorridos, o medias muestrales y desviaciones típicas, según se estime la desviación típica del proceso a partir del recorrido medio o a partir de las desviaciones típicas muestrales.

En cualquiera de los dos casos, el gráfico de medias muestrales estándar, utilizado con el fin de controlar la media del proceso, es equivalente tal y como se explica en el Anejo 1, a un contraste de hipótesis sobre la media del proceso. Por tanto, tiene sentido hablar de los conceptos de riesgo de primera y segunda especie en su acepción habitual.

La cuestión estadística básica a considerar en la fase de diseño, previa a la implantación de un gráfico de medias, es la elección de tres parámetros, de los que dependen la potencia del gráfico frente a distintas situaciones de falta de control así



como su comportamiento bajo control. Estos parámetros son el tamaño de muestra  $n$ , la frecuencia de muestreo  $f$ , definida como el intervalo de tiempo entre dos muestras sucesivas, y la posición de los límites de control  $m \pm L\sigma/\sqrt{n}$ .

El riesgo de primera especie, o porcentaje de falsas alarmas, depende sobre todo en el procedimiento estándar de la posición de los límites de control. Shewhart propone el valor  $L=3$ , asumiendo que una vez fijados los valores  $n$  y  $f$ , según otro tipo de consideraciones, y suponiendo que el estadístico a registrar sigue una distribución aproximadamente normal, el gráfico resultante presenta un equilibrio bastante razonable entre el riesgo de primera especie ( $\alpha = 0.0027$ ), y el de segunda  $\beta$ . El valor de este último depende a su vez del tamaño de muestra y de la magnitud de desviación de la media. En cuanto al tamaño de muestra, propone como regla práctica un  $n$  entre 4 y 6, siendo el valor 5 el más frecuentemente utilizado (al menos en la industria española).

En la bibliografía consultada sobre el tema, se pueden encontrar otras recomendaciones prácticas para los valores de  $n$  y  $f$ . Entre otras, Ishikawa (1976) plantea un ejemplo, en el que el plan de muestreo consiste en tomar muestras de tamaño  $n=5$  cada  $f=8$  horas. Por otro lado, Feigenbaum (1961) sugiere  $n=5$  y  $f=1$  hora, es decir muestras mucho más frecuentes que en el caso anterior. Sin embargo, en ninguna de estas referencias se realiza un estudio comparativo con otros planes alternativos, en función del coste de muestreo, o incluso en función de su rapidez para detectar una determinada desviación en la media del proceso.

En general se recomienda (Juran et al., 1990; Burr, 1953) utilizar tamaños de muestra de 2 a 6, sobre todo si la desviación típica del proceso  $\sigma$  se va a estimar a partir del recorrido. Para tamaños mayores, el recorrido medio dividido por la constante correspondiente,  $\bar{R}/d_2$ , resulta un estimador de  $\sigma$  sensiblemente inferior desde el punto de vista estadístico que  $s$ . Grant y Leavenworth (1985) señalan que para muestras de tamaño mayor que 15 no es aconsejable utilizar el recorrido para la estimación de la desviación típica  $\sigma$ . Teniendo en cuenta por otro lado la mayor simplicidad de cálculo

del gráfico de medias y recorridos respecto al de medias y desviaciones típicas, conviene en la práctica utilizar tamaños de muestra no excesivamente grandes.

La mayor parte de las referencias consultadas sobre el tema coinciden en señalar que tampoco interesa utilizar tamaños pequeños, cuando no se puede asumir distribución normal para la característica de calidad, ya que entonces tampoco se puede considerar distribución aproximadamente normal para la media muestral. Sin embargo con tamaños superiores a 3, y como resultado del Teorema Central del Límite, se puede admitir con menor margen de error normalidad en el estadístico a controlar.

Burr (1967) analiza el efecto que tiene sobre el gráfico de medias muestrales el incumplimiento de la hipótesis de normalidad en la característica de calidad, concluyendo que sólo implica algún problema cuando la distribución de probabilidad de la característica en cuestión es extremadamente asimétrica. Otros estudios empíricos (Ford, 1990) apoyan esta conclusión.

Además de las recomendaciones prácticas arriba consideradas, se encuentran también algunos trabajos que estudian con mayor profundidad el diseño del gráfico de medias, utilizando para la elección de los parámetros la función de potencia del gráfico, la Longitud de Recorrido Medio (ARL), o bien criterios económicos, como coste mínimo de operación. Los dos primeros son equivalentes ya que el ARL, como se ve más adelante, está estrechamente relacionado con la potencia del gráfico.

Dentro del primer grupo, se encuentra el trabajo de Knappenberger (1969), que se basa en la función de potencia para determinar los valores de  $n$  y  $L$ . Asume de partida que la característica a controlar sigue una distribución normal de media  $m$  y desviación típica  $\sigma$ . El procedimiento de diseño propuesto en su trabajo parte de la equivalencia existente entre el gráfico de medias y el contraste de hipótesis entre:

$$H_0: m = m_0$$

$$H_1: m = m_0 \pm \delta\sigma$$

donde  $\delta\sigma$  representa la magnitud de desviación de la media del proceso, y  $m_0$  la posición del mismo estando bajo control, que puede ser un valor estándar o bien

estimado como la media de las medias muestrales, a partir de un determinado número de subgrupos iniciales. Teniendo esto en cuenta, la probabilidad de que un punto caiga fuera de los límites de control cuando el proceso está bajo control es igual a  $\alpha$ , o riesgo de primera especie:

$$P(\bar{X} \geq m_0 + L\sigma/\sqrt{n} / m=m_0) + P(\bar{X} \leq m_0 - L\sigma/\sqrt{n} / m=m_0) = \alpha$$

o lo que es lo mismo

$$P(z \geq L / m=m_0) = \alpha/2 \quad (1)$$

con lo que,  $L=Z_{\alpha/2}$ , siendo  $Z_t$  el valor de una normal tipificada que tiene una probabilidad de ser superada igual a  $t$ . La potencia de este test vendrá dada por la expresión:

$$P(\bar{X} \geq m_0 + Z_{\alpha/2} \sigma / \sqrt{n} / m=m_0 \pm \delta\sigma) + P(\bar{X} \leq m_0 - Z_{\alpha/2} \sigma / \sqrt{n} / m=m_0 \pm \delta\sigma) = (1-\beta)$$

Si los límites de control se sitúan a dos o tres veces la desviación típica muestral respecto de la media, y el orden de magnitud de la desviación que cabe esperar es aproximadamente igual a  $\sigma$ , la ecuación anterior se simplifica a:

$$Z_{\alpha/2} - \delta \sqrt{n} = Z_{1-\beta}$$

de donde:

$$n = (Z_{\alpha/2} - Z_{1-\beta})^2 / \delta^2 \quad (2)$$

Una vez fijados  $\alpha$ ,  $\beta$  y  $\delta$ , el valor de  $L$  vendrá dado por la expresión (1) y el de  $n$  por la (2).

Por lo tanto, en esta aproximación los parámetros de diseño del gráfico de medias se eligen en función de los valores  $\alpha$  y  $\beta$ , sin tener en cuenta los costes asociados al procedimiento resultante. La simplificación asumida respecto a un orden de magnitud de desviación importante igual a  $\sigma$ , puede llevar a su vez a la elección de unos parámetros que no se correspondan con lo que realmente va a pasar en el proceso, pues no siempre se conoce a priori el valor de la desviación que tendrá lugar en la media. Por otro lado, la frecuencia de muestreo  $f$  queda sin definir, ya que no se contempla su elección en el procedimiento anterior.

El valor esperado del número de artículos defectuosos producidos entre la salida de control y su detección en un gráfico es también una medida de la efectividad del mismo, cuando realmente hay problemas en el proceso, así como del número de falsas alarmas estando bajo control. Aroian y Levene (1950) introducen este parámetro, definiendo lo que denominan AEN (Average Efficiency Number) y ASSN (Average Stoppage Spacing Number).

Posteriormente, Page (1954) define el ARL (Average Run Length) o Longitud de Recorrido Medio, como el valor esperado del número de muestras de tamaño  $n$  inspeccionadas desde que se produce la salida de control en el proceso hasta que se detecta. Este parámetro es denominado también ASRL (Average Sample Run Length). Si el tamaño de muestra es constante, el ARL está a su vez directamente relacionado con el AARL (Average Article Run Length), que es el número de artículos a muestrear desde que se produce la falta de control hasta que el gráfico lo detecta. Estos dos parámetros se utilizan habitualmente para el diseño, entre otros, del gráfico de medias.

Si se desea detectar desviaciones positivas y negativas en la media del proceso, asumiendo una distribución normal, se tiene que los valores del AARL estando bajo control  $L_0$  y fuera de control  $L_1$ , vienen dados por:

$$L_0 = n/2[1 - \phi(L)]$$

$$L_1 = n/[1 - \Phi(L - \delta\sqrt{n}) + \Phi(-L - \delta\sqrt{n})]$$

donde

$$\Phi(x) = \int_{-\infty}^x (2\pi)^{-1/2} \exp\left(-\frac{x^2}{2}\right) dx$$

El método propuesto por Page consiste en fijar el valor de  $L_0$ , y dado  $\delta$  elegir los parámetros  $L$  y  $n$  que minimizan  $L_1$ , o bien especificar  $L_1$  y elegir  $L$  y  $n$  maximizando  $L_0$ . Como se verá en el apartado de metodología, el valor de  $L_0$  es igual a  $n/2\alpha$ , siendo  $\alpha$  la probabilidad de que aparezca un punto fuera de uno de los límites de control, o riesgo de primera especie.

Al igual que ocurría antes, el procedimiento de diseño propuesto implica conocer la magnitud  $\delta$  de desviación cuya detección precoz resulta relevante, y adicionalmente no permite fijar el valor del intervalo de muestreo  $f$ . Por otro lado, no tiene en cuenta ninguna consideración de tipo económico a la hora de diseñar el gráfico de medias.

En la misma línea se encuentra la propuesta de Weiler (1952). Asumiendo una desviación positiva  $\delta$  en la media del proceso, y fijando los límites en una posición  $L$  determinada, el autor encuentra que el tamaño de muestra  $n$  que minimiza el AARL, o valor esperado del número de artículos a muestrear hasta que se detecta esta desviación, es función de la magnitud de estos dos parámetros. Por ejemplo, si  $L=3.09$ , el valor de  $n$  resulta igual a  $12.0/\delta^2$ , pero cuando  $L=3$ ,  $n$  es igual a  $11.1/\delta^2$ , etc. Ahora bien, el método propuesto por Weiler no es más que un caso particular del de Page, siendo más restrictivo y presentando los inconvenientes señaladas antes para este último.

Finalmente, algunos trabajos abordan el diseño del gráfico de medias desde el punto de vista económico, tratando de minimizar el coste de operación del mismo. El primero en considerar este enfoque es Duncan (1956), que formula un modelo de costes asumiendo ciertas simplificaciones en el procedimiento de control y muestreo. Sin embargo, la ecuación que se deriva permite calcular los valores óptimos de  $n$ ,  $f$  y  $L$  sólo mediante tanteo.

Goel (1968) y, posteriormente, Chiu et al.(1974), Montgomery (1980) y Lorenzen et al. (1986) abordan la resolución analítica del modelo propuesto por Duncan, introduciendo algunas modificaciones. De todos ellos, el modelo establecido por Lorenzen es el más general, y permite calcular el valor esperado del coste horario de operación del gráfico de medias, proponiendo un algoritmo que emplea de forma iterativa tres procedimientos de minimización, partiendo de un valor inicial para el tamaño de muestra  $n=20/\delta$  y para  $L$  de forma que  $L_1= 1.2$ . Una vez aplicado este algoritmo, el resultado obtenido es el coste mínimo y los parámetros de diseño  $n$ ,  $f$  y  $L$ , correspondientes al mismo.

Los principales inconvenientes de esta aproximación al problema de diseño del gráfico de medias son la enorme complejidad del procedimiento de cálculo, y la dificultad que conlleva determinar la gran cantidad de parámetros que se necesita conocer para poder elaborar el modelo. La ecuación propuesta depende de 12 parámetros de coste y tiempo del proceso, 2 variables de estado que indican si la producción continúa durante la búsqueda y eliminación de causas especiales de variabilidad, y de los tres parámetros de diseño  $n$ ,  $f$  y  $L$  del gráfico.

A excepción de estos tres últimos, los otros deben estar completamente determinados antes de aplicar el algoritmo correspondiente; un análisis de sensibilidad posterior muestra que la solución encontrada puede verse modificada por pequeñas variaciones en estos valores, dependiendo también del valor inicial elegido para  $n$ . Adicionalmente, los autores indican que el intervalo de muestreo  $f$ , salvo excepciones, resulta ser un valor poco conveniente a nivel práctico.

Por otro lado, Woodall (1986) señala los inconvenientes de este método, ya que en general conduce al diseño de procedimientos que presentan unas propiedades estadísticas poco convenientes, pues aumenta la cantidad de falsas alarmas y disminuye la sensibilidad en la detección de ciertas desviaciones. Tal y como indica este autor, en el modelo asumido por Duncan se tienen en cuenta sólo los costes a corto plazo y los debidos a cantidad de producto fuera de especificaciones. Esto entra en contradicción con la moderna filosofía de control de calidad, en la que uno de los objetivos es asegurar la supervivencia de la empresa en el mercado a largo plazo, y en la que se asume como principio básico la necesidad de minimizar la variabilidad del proceso centrado en el nominal.

Adicionalmente, los procedimientos complicados para el diseño del gráfico y plan de muestreo no resultan una solución adecuada a este problema. Se debe tener en cuenta que los gráficos de control constituyen una herramienta útil para la resolución de problemas on-line, entre otras cosas, por su sencillez y facilidad de manejo en planta.

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 del mismo, 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 algunas situaciones, frecuentes en la práctica, el coste de muestreo resulta aproximadamente proporcional al número de unidades muestreadas. Para una misma cantidad total de artículos seleccionados, se plantea en estos casos el dilema de elegir entre planes de muestreo consistentes en muestras pequeñas a intervalos cortos de tiempo, y muestras más grandes pero menos frecuentes. La diferencia entre uno u otro plan estará en la cantidad de artículos defectuosos producidos antes de que el gráfico detecte la salida de control del proceso. No habiendo encontrado respuesta a esta cuestión en la bibliografía consultada sobre el tema, nos planteamos abordarla según lo que se explicita en el apartado de Objetivos.

### **1.2.2- Señales adicionales de falta de control en gráficos de medias**

El gráfico de medias muestrales es uno de los procedimientos más extendidos para control on-line de procesos en los que la calidad viene definida por el valor de una característica medible en una escala continua. Una vez abordado el diseño óptimo de este gráfico, tiene interés el estudio de sus propiedades, con el fin de determinar su efectividad frente a las distintas situaciones de falta de control que pueden darse en el proceso.

La primera característica a tener en cuenta es que el gráfico de medias diseñado de forma estándar, con límites de control situados a  $m_0 \pm 3\sigma/\sqrt{n}$ , resulta un procedimiento poco efectivo en la detección de desviaciones de la media del proceso de poca magnitud. Existen tres posibles cursos de acción que pueden, hasta cierto punto, mejorar el gráfico en este sentido.

La primera posibilidad consiste en adoptar un tamaño de muestra grande, ya que asumiendo que el intervalo de muestreo no es un parámetro de diseño crítico, esta medida aumentará en general la potencia del gráfico. El tamaño a utilizar será tanto mayor cuanto menor sea la magnitud de desviación cuya detección precoz es relevante. Sin embargo, la mejora en sensibilidad que se consigue con este método puede ser insuficiente y no compensar el mayor coste asociado a la toma de muestras mayores.

La segunda alternativa consiste en aumentar la potencia a costa de incrementar también el riesgo de primera especie, utilizando unos límites de control más estrechos. Convendrá recurrir a este método en función del coste asociado a ajustar el proceso cuando realmente no hace falta, ya que el porcentaje de falsas alarmas obviamente aumenta en este caso.

La tercera posibilidad, y la más utilizada en la práctica, consiste en aprovechar la información acumulada en el gráfico de medias a base de introducir señales de falta de control adicionales, además de la propuesta por Shewhart (punto fuera de los límites de control). Esta medida se basa en considerar que cuando se produce una desviación en la media, su efecto se manifiesta en varias muestras antes de que aparezca algún punto fuera de los límites. Una manera de aprovechar esta información consiste en definir señales de falta de control, que estén basadas en pautas o rachas que pueden aparecer en las distintas zonas en las que se divide el gráfico de medias. Un esquema y definición de estas zonas se recoge en el Anejo 1.

Se puede encontrar en la bibliografía sobre el tema, sugerencias sobre señales adicionales de falta de control. Dudding & Jennett (1944) proponen la utilización de límites de aviso  $m \pm 2\sigma/\sqrt{n}$  en el gráfico de medias, de forma que se considera señal de falta de control si un cierto número de puntos aparece entre estos límites y los de control (Zona A), situados en la posición estándar. Page (1955 y 1964) estudia las propiedades estadísticas del gráfico de medias con límites de aviso, introduciendo como señal de falta de control adicional la aparición de  $k$  puntos sucesivos en la Zona A, en los casos de  $k=3$  y  $4$ .



La primera propuesta formal de señales adicionales se puede encontrar en Western Electric Handbook (1956), y a partir de estas, Ishikawa (1974), Bissell (1978), Wheeler (1983), Ford (1984), Champ et al. (1989) y Walker et al. (1991), entre otros, sugieren otras pautas o tendencias en el gráfico de medias que pueden ser también indicio de falta de control.

En general, todas estas señales pueden resumirse básicamente de la siguiente forma: tiene lugar una señal de falta de control cuando  $k$  de las  $p$  últimas medias muestrales forman una racha de determinadas características. Según esto, no se considera sólo la posición del último punto graficado, sino también la información contenida en un determinado número de muestras anteriores.

En el contexto de la industria española, las señales de falta de control utilizadas más frecuentemente ( Ford,1984 ) son las siguientes (Ishikawa, 1974):

**Señal 1:** Punto fuera de los límites  $m \pm 3 \sigma/\sqrt{n}$  .

**Señal 2:** Dos de tres puntos seguidos en la zona comprendida entre los límites de aviso  $m \pm 2 \sigma/\sqrt{n}$  y los de control  $m \pm 3 \sigma/\sqrt{n}$  .

**Señal 3:** Siete puntos seguidos por encima o por debajo de la línea central.

**Señal 4:** Racha ascendente o descendente de siete puntos seguidos.

Las tres señales adicionales anteriores son similares aunque más restrictivas que algunas de las definidas en Western Electric Handbook. Por ejemplo, la racha correspondiente a la Señal 3 tiene una longitud mínima de ocho puntos seguidos.

El introducir nuevas señales de falta de control en el gráfico de medias supone, en primer lugar, una complejidad de manejo mucho mayor respecto del procedimiento estándar propuesto por Shewhart. A nivel práctico tiene interés estudiar si la mejora que se produce en sensibilidad compensa la dificultad que puede suponer su utilización en planta. Por otro lado, las propiedades estadísticas del gráfico cambian, aumentando al mismo tiempo que la sensibilidad a bajas desviaciones el porcentaje de falsas alarmas estando bajo control.

Aunque las señales de falta de control adicionales se han venido utilizando durante muchos años, no existe ningún estudio sobre el comportamiento y propiedades del gráfico de medias cuando se utilizan las definidas antes, si bien Page (1962), Wheeler (1983), Champ et al. (1987) y Walker (1991) abordan el estudio de las propiedades estadísticas del gráfico resultante de introducir algunas combinaciones de señales adicionales distintas a la mencionada antes.

El cálculo analítico del ARL correspondiente a un gráfico de medias con señales adicionales resulta una tarea extremadamente compleja, e incluso inabordable para alguna combinación concreta. Se puede demostrar fácilmente que el ARL en el gráfico de medias estándar es el inverso de la probabilidad de que se presente señal de falta de control. En consecuencia, el ARL bajo control será  $1/\alpha$ , siendo  $\alpha$  el riesgo de 1ª especie o probabilidad de falsa alarma.

Cuando se incluyen  $t$  señales de falta de control independientes entre sí, Duncan (1986) propone la siguiente expresión para  $\alpha$  :

$$\alpha = 1 - \prod_{i=1}^t (1 - \alpha_i)$$

en la que  $\alpha_i$  es el riesgo de primera especie correspondiente a la señal  $i$ .

Sin embargo, esta aproximación no resulta válida si se consideran las cuatro señales antes definidas. En primer lugar, estas señales no son independientes entre sí. Adicionalmente, la definición dada para  $\alpha_i$  no es exactamente correcta, ya que el riesgo de primera especie asociado a la señal  $i$  no es un valor constante, puesto que depende de observaciones pasadas. Por tanto, la expresión anterior no es aplicable a este caso.

Cuando la falta de control puede deberse a señales que dependen de parte de la historia pasada del gráfico, la semejanza con el test de hipótesis clásico para la media desaparece. La definición dada para el valor de  $\alpha$  como "probabilidad de falsa alarma" puede llevar a confusión, ya que no está claro si se refiere a la probabilidad de falsa

alarma en un instante concreto, magnitud que varía según observaciones pasadas, o a la frecuencia relativa de falsas alarmas.

Definiendo  $\alpha$  como la frecuencia relativa de falsas alarmas, la relación ARL bajo control =  $1/\alpha$  es directa. Por lo tanto, resulta más conveniente en este caso estudiar la eficiencia del gráfico a partir de los valores del ARL en distintas situaciones de falta de control.

Page (1962) obtiene valores del ARL para el procedimiento que propone, aproximando la distribución de la longitud del recorrido (run length) a la del número de transiciones en una cadena de Markov antes de alcanzar un estado determinado, en este caso la señal de falta de control debida a un punto fuera de los límites de control o a  $k$  ( 3 ó 4 ) puntos entre estos y los de aviso. En general, el ARL de un gráfico de medias con estas señales, diseñado para detectar desviaciones positivas y negativas, puede obtenerse a partir de  $ARL_s$  y  $ARL_i$ , correspondientes a los procedimientos diseñados para detectar desviaciones de la media en un solo sentido, utilizando la expresión (Kemp,1961):

$$1/ARL = 1/ARL_s + 1/ARL_i$$

Normalmente los límites de control se sitúan en el gráfico de medias en  $m \pm 3\sigma/\sqrt{n}$ , y los de aviso en  $m \pm 2\sigma/\sqrt{n}$ , de forma que si la media del proceso no ha sufrido ninguna desviación,  $ARL_s=ARL_i$ , con lo que el ARL bajo control se puede calcular como  $ARL_s/2$ . Cuando la media sufre una desviación igual a  $\delta\sigma$ , Page deduce a partir de la aproximación que asume, las siguientes expresiones para  $ARL_s$  y  $ARL_i$ :

$$ARL_s = \frac{1 - r^n}{1 - r - p(1 - r^n)}$$

$$ARL_i = \frac{1 - s^n}{1 - s - q(1 - s^n)}$$

donde  $p$  y  $r$  son respectivamente las probabilidades de que un punto caiga por debajo del límite de aviso superior, y entre este límite y el de control superior.  $q$  y  $s$  son a su vez las probabilidades de que un punto caiga por encima del límite de aviso inferior, y entre

éste y el límite de control inferior. Estos parámetros se pueden calcular, asumiendo distribución normal para la característica a controlar, de la siguiente manera:

$$p = \Phi(2 - \delta \sqrt{n}), \quad q = 1 - \Phi(-2 - \delta \sqrt{n})$$

$$r = \Phi(3 - \delta \sqrt{n}) - \Phi(2 - \delta \sqrt{n})$$

$$s = \Phi(-2 - \delta \sqrt{n}) - \Phi(-2 - \delta \sqrt{n})$$

donde

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

Utilizando las expresiones anteriores, se calcula el valor del ARL para  $k=3$  y  $4$ , y tamaños de muestra  $n=5$  y  $10$ , aunque no se comparan los procedimientos resultantes con el gráfico de medias estándar con límites de control ajustados para equiparar el porcentaje de falsas alarmas. Queda sin resolver, por tanto, la cuestión relativa a la mejora de efectividad del gráfico teniendo en cuenta además que el procedimiento de cálculo propuesto no es válido cuando se consideran conjuntamente las cuatro señales antes definidas.

Wheeler (1983) calcula la potencia del gráfico de medias, diseñado para detectar desviaciones de la media en un solo sentido, tomando además de la señal basada en un punto fuera de los límites de control, tres señales adicionales propuestas en Western Electric Handbook (1956). Ninguna de ellas tiene relación con la Señal 4, racha ascendente o descendente de 7 puntos seguidos. Por otro lado, las tablas que aparecen en este artículo solamente reflejan la probabilidad de que se presente alguna de las señales asumidas en  $k$  subgrupos, variando  $k$  de 1 a 10. Los valores se han obtenido analíticamente, para  $k$  pequeño, y mediante simulación de Monte Carlo para  $k$  grande.

Walker et al. (1991) amplían el estudio realizado por Wheeler, calculando mediante simulación la probabilidad de falsa alarma en  $k=20$  y  $30$  subgrupos, tomando además de las señales de falta de control consideradas por este último, una señal de racha pero de longitud seis.

Los resultados de estos dos trabajos ponen de manifiesto que la utilización de señales adicionales incrementa el porcentaje de falsas alarmas en el gráfico de medias, cuando se considera un determinado número de subgrupos  $k$ . Sin embargo, ninguno de los dos compara la efectividad del gráfico resultante con la de otros procedimientos alterativos, como el gráfico de medias con una sola señal pero límites más estrechos, o el gráfico CUSUM. Adicionalmente, en ambos casos la desviación asumida en la media del proceso es constante, no contemplándose la posibilidad de desviación progresiva en la misma. (En el caso de desviación constante de magnitud  $\delta\sigma$ , la media del proceso se mantiene fija a un nuevo nivel  $m+\delta\sigma$  y en el caso de desviación progresiva la nueva media es igual a  $m+\delta i\sigma$ , siendo  $i=1,2,\dots$ , el número de muestra desde que el proceso se salió de control).

Champ y Woodall (1987) abordan el cálculo del ARL para distintas combinaciones de rachas o señales de falta de control según el procedimiento propuesto por Page, es decir, aproximando el gráfico con señales adicionales a una cadena de Markov. A diferencia del trabajo de Page, introducen combinaciones de hasta seis señales diferentes, entre las que no se incluye la señal de racha ascendente o descendente. También en este caso se consideran sólo desviaciones constantes en la media.

El procedimiento utilizado en estos trabajos resulta sin embargo muy complicado por el elevado número de estados que requiere para modelizar el gráfico cuando aumenta la cantidad de señales adicionales introducidas al mismo tiempo, y no es aplicable cuando se considera la Señal 4, racha ascendente o descendente de varios puntos.

En el caso de deriva progresiva, existen muy pocas referencias en las que analice el comportamiento y propiedades de los procedimientos más habituales. Bissell (1984) aborda el estudio comparativo de la efectividad del gráfico CUSUM, gráfico de medias estándar y gráfico de medias con límites de aviso, aproximando los gráficos a una cadena de Markov no homogénea. Los resultados del estudio son discutibles puesto que

no se consideran las señales basadas en rachas y además los gráficos a comparar no se ajustan para equiparar sus ARL bajo control.

Davis y Woodall (1988) también estudian, pero recurriendo en este caso a técnicas de simulación, la efectividad del gráfico de medias cuando se considera un conjunto de cuatro señales, además de la racha ascendente o descendente. Sin embargo, no se aborda la comparación de los procedimientos resultantes con otros alternativos, con el fin de determinar si se produce una mejora en su eficiencia en la detección de desviaciones de poca magnitud.

### **1.2.3- Gráficos CUSUM**

Los gráficos de sumas acumuladas, o gráficos CUSUM, fueron introducidos por Page en 1954, y a partir de entonces se han propuesto diversas modificaciones, existiendo una amplia bibliografía sobre su diseño óptimo y propiedades. En el Anejo 1, puede encontrarse una descripción de este procedimiento, basada en algunas de estas referencias.

Una encuesta realizada recientemente por Peña et al.(1990), sobre la utilización de métodos estadísticos por las empresas de la A.E.C.C., pone de manifiesto que las técnicas de control de procesos CUSUM son utilizadas por un porcentaje de empresas españolas muy pequeño. Un estudio similar en la American Society for Quality Control (Saniga y Shirland, 1977), refleja que ya en la década de los setenta, el porcentaje de empresas americanas que utilizaban estos procedimientos, aunque bajo, era superior al registrado en la encuesta antes mencionada.

Por otro lado, Lucas (1991) señala que en determinadas empresas americanas del sector químico, los gráficos CUSUM constituyen hoy en día la herramienta habitual para control on-line de procesos.

El motivo de su relativamente escasa implantación, con respecto a los gráficos propuestos por Shewhart, puede encontrarse en la mayor dificultad y complejidad de

diseño del gráfico CUSUM. Ahora bien, una vez diseñado, su manejo puede ser tanto o más sencillo que, por ejemplo, el del gráfico de medias implementado con señales adicionales de falta de control (Lucas, 1976).

Por otro lado, existen dos versiones del gráfico CUSUM, completamente equivalentes entre sí, desde el punto de vista de su efectividad, pero que pueden diferir en complejidad de manejo. Aunque cualquiera de las dos admite representación gráfica, la relacionada con el valor de referencia  $k$  y el intervalo de decisión  $h$ , puede implementarse sólo numéricamente, tabulando el valor de dos sumas acumuladas. La descripción de estas dos versiones, ilustrada con un ejemplo, se encuentra en el Anejo 1.

Los gráficos CUSUM, tal y como se manejan hoy en día, se derivan de los propuestos por Page (1954). La idea de partida es crear un procedimiento de control secuencial que, a diferencia del gráfico de medias estándar, aproveche la información contenida en todas las observaciones. El método para detectar desviaciones de la media en un solo sentido se basa en el cálculo de la suma acumulada de puntuaciones asignadas según un artículo sea defectuoso o no, y se considera señal de falta de control el que esta suma acumulada supere una cierta cantidad  $h$ .

Barnard (1959) introduce la versión consistente en graficar la suma acumulada de las desviaciones de las observaciones respecto de la media, y comprobar si el proceso está bajo control o no utilizando una máscara en  $V$ . Proporciona además una descripción de los dos parámetros de la máscara,  $d$  distancia entre el último punto graficado y el vértice, y  $\Theta$  ángulo entre la horizontal y cualquiera de las dos ramas de la máscara. Tendrá lugar una señal de falta cuando el valor de la suma acumulada caiga fuera de alguna de las ramas de la máscara en  $V$ .

Posteriormente, Ewan y Kemp (1960) recogen la propuesta de Page, pero introduciendo los parámetros  $k$  valor de referencia y  $h$  intervalo de decisión. Adicionalmente deducen la relación existente entre éstos y los parámetros  $d$  y  $\Theta$  de la máscara en  $V$ , mostrando la equivalencia entre ambos procedimientos.

Además de estos tres artículos, importantes puesto que definen un nuevo procedimiento para control on-line de procesos, existen numerosas referencias sobre los CUSUM, que pueden agruparse según el aspecto de los mismos que abordan. Así, un primer grupo de artículos trata fundamentalmente sobre su utilización y la elección óptima de sus parámetros. Algunos de estos trabajos analizan el cálculo del ARL del CUSUM para control de la media. Finalmente, cabe considerar una serie de artículos sobre modificaciones a introducir en el procedimiento CUSUM estándar, con el objetivo de mejorar su efectividad en algún sentido, tal y como se discute más adelante.

Con respecto al problema de diseño o elección óptima de sus parámetros, Barnard (1959) propone inicialmente fijar los valores  $d$  y  $\Theta$  mediante tanteo. Un método de diseño de gráficos CUSUM bastante extendido, es el propuesto por Johnson (1961), basado en la aproximación de estos gráficos a un test secuencial de ratio de verosimilitudes. Vease por ejemplo Johnson y Leone (1977), Wadsworth et al. (1986), SAS Institute Inc. (1986), Statgraphics (1989), entre otros. En este método se especifica la máxima probabilidad deseable de falsa alarma y la probabilidad de no detectar una desviación de magnitud  $\delta$ . Estas dos probabilidades se representan respectivamente como  $2\alpha$  y  $\beta$ . Johnson recomienda calcular los parámetros de la máscara a partir de:

$$\Theta = \arctan (\delta/2)$$

y

$$d = 2\delta^{-2} \ln[(1-\beta)/\alpha]$$

tomando factor de escala 1 (1 Ud de la escala horizontal corresponde a 1 Ud de la vertical ). Johnson (1961) señala que debido a que el valor de  $\beta$  es normalmente muy pequeño, se puede simplificar la última ecuación a:

$$d \approx -2\delta^{-2} \ln (\alpha)$$

Aplicando, la equivalencia deducida por Ewan y Kemp (1960), se tiene que los parámetros de la otra posible versión de CUSUM son:

$$\text{valor de referencia } k = \delta/2$$

$$\text{intervalo de decisión } h \approx -\ln(\alpha)/\delta$$



Adams, Lowry y Woodall (1991) señalan que la definición de probabilidad de falsa alarma dada para el valor  $2\alpha$  es ambigua y que el parámetro  $\beta$  no tiene una clara interpretación para el gráfico CUSUM. La probabilidad de falsa alarma en un gráfico de este tipo varía con cada nuevo punto. La única interpretación con sentido para el valor  $2\alpha$  es el de proporción a largo plazo de observaciones que resultan en falsa alarma. El ARL bajo control correspondiente es igual entonces a  $1/(2\alpha)$ . A pesar de esta consideración, el método de Johnson resulta muy inexacto. Woodall y Adams (1991) calculan los ARL bajo control de gráficos CUSUM diseñados según este método, y muestran que son aproximadamente cinco veces el valor que cabría esperar de la aproximación basada en la relación  $ARL = 1/(2\alpha)$ .

Según el método de Johnson, un gráfico CUSUM diseñado para un  $\alpha=0.00135$ , presenta un porcentaje de falsas alarmas comparable al de un gráfico de medias de Shewhart con límites 3-sigma. Sin embargo, para este último el ARL bajo control resulta igual a 370.4, mientras que a partir de los resultados proporcionados por Woodall y Adams, se observa que los ARL de los gráficos CUSUM correspondientes, cuyos parámetros se han calculado a partir de las fórmulas de Johnson, están cercanos a 2000.

Por lo tanto, utilizando el método de Johnson se obtienen gráficos CUSUM con un porcentaje de falsas alarmas mucho menor de lo que se pretende, y esto a su vez implica una efectividad más baja en la detección de desviaciones en la media. No es posible ajustar los parámetros del gráfico CUSUM con el fin de obtener un porcentaje de falsas alarmas bajo sin modificar su potencia.

En general, se recomienda diseñar los gráficos CUSUM en función del ARL bajo control y de la desviación cuya detección precoz resulta relevante. Kemp (1961) discute la elección de los parámetros  $k$  y  $h$  del CUSUM, basándose en estos criterios. En otro trabajo, el mismo autor (Kemp, 1962) proporciona un nomograma a partir del cual es posible abordar el diseño de una forma relativamente sencilla. En esta última referencia, el autor propone la siguiente regla práctica: si la magnitud de desviación que se desea detectar es igual a  $\delta\sigma/\sqrt{n}$ , dado un ARL para esta desviación, el CUSUM con

un valor de referencia  $k=\delta/2$  maximiza el ARL bajo control (es decir, minimiza el porcentaje de falsas alarmas ).

Esta última recomendación es aceptada en todos los trabajos que hasta la fecha han tratado el problema de diseño óptimo del gráfico CUSUM. Por citar uno de los más recientes, Woodall y Adams (1991) recomiendan el siguiente procedimiento: Si se desea especificar la proporción de falsas alarmas,  $2\alpha$ , entonces el ARL bajo control es igual a  $1/(2\alpha)$  y dado un valor crítico para la desviación  $\delta$ , tomar  $k=\delta/2$  y a continuación elegir el  $h$  correspondiente al ARL bajo control anteriormente fijado. Si se va a implantar la versión del CUSUM con máscara en V, los parámetros  $d$  y  $\Theta$  se obtienen a partir de  $k$  y  $h$  utilizando las expresiones proporcionadas por Ewan y Kemp (1960).

Una buena discusión sobre la utilización e implementación de estos gráficos se puede encontrar en Ewan (1963), donde se compara además su eficiencia con alguno de los procedimientos estándar. Otras referencias sobre este tema son Van Dobben de Bruyn (1968), Bissell (1969) y Goel y Wu (1971), encontrándose en este último trabajo los nomogramas más exactos para el diseño de gráficos CUSUM para la media de una característica con distribución normal.

Lucas (1976) señala que de las dos versiones del CUSUM, la que maneja el valor de referencia  $k$  y el intervalo de decisión  $h$ , resulta, tanto en su versión tabulada como gráfica, más fácil de implementar que la máscara en V. Como inconvenientes de la segunda señala la imposibilidad de diseñar un impreso en el que llevar el gráfico, puesto que la suma acumulada puede tomar valores muy dispares, y por otro lado la no existencia de un criterio sobre el número de puntos a considerar en cada instante al aplicar la máscara.

Sin embargo, cuando se utiliza la máscara en V, si se ha elegido un factor de escala adecuado, se detecta visualmente el punto en que ha tenido lugar una cierta desviación en la media más rápidamente que la otra versión, incluso antes de que la suma acumulada supere los límites de la máscara. Un ejemplo de la ventaja "visual" del

gráfico CUSUM se puede encontrar en la aplicación desarrollada por Romero et al. (1987) para control de consumo de gas en una industria de automóviles.

Ewan (1963) propone algunas simplificaciones en el diseño de la máscara en V. Recomienda construir máscaras que puedan ser utilizadas con distintos gráficos CUSUM, siempre que sea posible, y define como procedimiento CUSUM estándar el diseñado con un  $k=0.5\sigma/\sqrt{n}$  y un  $h=5\sigma/\sqrt{n}$ , que presenta un ARL bajo control alrededor de 500 y una efectividad en general buena.

Con respecto al factor de escala, sugiere seguir la recomendación de Barnard (1959), que propone un valor igual a  $2\sigma/\sqrt{n}$ , es decir 1 Ud de escala horizontal equivale a  $2\sigma/\sqrt{n}$  Uds. de escala vertical. Un factor de escala superior supondría una pérdida de capacidad para detectar visualmente desviaciones en la media del proceso. Hay que tener en cuenta que este factor determina la magnitud de desviación que provoca un cambio de pendiente en el gráfico de  $0^\circ$  a  $45^\circ$ .

Por otro lado, también señala que, en general, resulta difícil la implantación adecuada de la máscara en V para control de procesos, siendo preferible recurrir a otros procedimientos, como por ejemplo el basado en k y h. Una aplicación de los gráficos CUSUM para control on-line en la industria química, sector en el que han encontrado mayor aceptación (Lucas, 1991), se encuentra en Butler (1983).

Como conclusión de lo anterior, podemos considerar que el diseño óptimo del gráfico CUSUM debe basarse en sus propiedades estadísticas, expresadas en función del ARL. Este parámetro fue introducido por Page (1954), tal y como se señaló al principio de esta revisión bibliográfica. El ARL es el valor esperado de la variable aleatoria número de muestras entre dos señales de falta de control consecutivas, denominada normalmente con el término longitud del recorrido (run length).

El cálculo del ARL del gráfico CUSUM ha sido objeto de discusión desde el momento en que se definió, existiendo numerosas referencias al respecto. En este caso resulta inabordable la determinación analítica de la distribución exacta de la variable

aleatoria longitud del recorrido (RL), por lo que para calcular su valor medio, ARL, se debe recurrir a métodos aproximados.

Por otro lado, Barnard (1959) señala que utilizar el ARL para caracterizar el comportamiento de un procedimiento CUSUM puede resultar engañoso, y recomienda dar más peso a los valores de la distribución de RL extremadamente grandes, proponiendo utilizar como medida de efectividad el valor esperado de esta variable al cuadrado.

Ewan y Kemp (1960) y Kemp (1962) discuten la propuesta de Barnard y consideran que si se utiliza el ARL para definir la efectividad de un CUSUM, tiene también sentido estudiar la determinación aproximada de la distribución de RL, así como otros parámetros de la misma (mediana, cuartiles, etc.). Ewan y Kemp (1960) deducen que cuando el valor del ARL es grande, la distribución de RL se asemeja a la geométrica, pero cuando ARL es pequeño, v.g. cuando la desviación de la media del proceso es muy grande, esta aproximación ya no es válida.

Se ha realizado una revisión bibliográfica de los distintos métodos propuestos para la determinación del ARL, porque en nuestra investigación hemos abordado el análisis del comportamiento del gráfico CUSUM en ciertos contextos, y ha sido necesario adoptar un procedimiento para el cálculo de sus propiedades.

Page (1954) plantea las ecuaciones básicas para cálculo del ARL, que resultan ser ecuaciones integrales de Fredholm de segunda especie. Su solución analítica puede encontrarse para ciertos casos particulares, por ejemplo, cuando las observaciones siguen la distribución Gamma. Sin embargo, en otros casos, incluyendo el más frecuente de distribución normal, se debe recurrir a técnicas de resolución numérica aproximada.

Ewan y Kemp (1960) aportan soluciones a este problema, a partir del test de Wald. Kemp (1962) proporciona un nomograma para el diseño del gráfico CUSUM, estimando los ARL mediante el mismo procedimiento.

El ábaco más exacto para diseño del gráfico CUSUM es el proporcionado por Goel y Wu (1971). En este caso, asumiendo distribución normal para la característica en cuestión, se aborda el cálculo del ARL aproximando las ecuaciones integrales mediante sistemas de ecuaciones lineales, según el método propuesto por Kantorovich y Krylov (1958). Estas ecuaciones a su vez son resueltas mediante cuadratura gaussiana.

Otro método para estimar este parámetro, es el propuesto por Brook y Evans (1972). En este caso se modeliza el gráfico CUSUM como una cadena de Markov, para los casos de variable aleatoria discreta y continua, pero asumiendo que la suma acumulada toma valores enteros. Con este modelo, calculan la matriz de transición y a partir de ésta el valor del ARL, estudiando también la distribución de RL.

Woodall (1983 y 1984) aborda la estimación de este parámetro, mejorando el método de Brook y Evans (1972). Aproxima el gráfico CUSUM a una cadena de Markov discreta, pero con un espacio de estados continuo y resuelve las ecuaciones derivadas aplicando el método de Simpson con un número variable de puntos de cuadratura, según la situación. Siguiendo la misma línea, Waldmann (1988) propone una modelización del gráfico CUSUM con un menor número de estados, y compara el método de Simpson con otros procedimientos de cuadratura numérica.

Woodall y Adams (1991) proponen a su vez un método alternativo basado en una aproximación de Siegmund (1985). Se trata de un procedimiento numérico, en el que se utiliza una ecuación de partida que exige fijar el ARL bajo control,  $k=\delta/2$ , y un parámetro  $b_0$ , con el que se itera según el método de Newton-Raphson, para encontrar el valor de  $h$  que corresponde al ARL elegido. La ventaja de este método es su simplicidad, ya que los cálculos que exige pueden ejecutarse con una calculadora de bolsillo. Por otro lado, los resultados que proporciona no son tan exactos como los de Goel y Wu (1972), y si el valor de  $b_0$  elegido para empezar no es el adecuado el número de iteraciones puede ser grande.

Dentro de los métodos numéricos, cabe considerar el propuesto por Reynolds (1975), similar a los anteriores en cuanto a exactitud y resultados, pero diferente pues deducen las ecuaciones correspondientes estableciendo una analogía entre la suma acumulada y el movimiento browniano.

Finalmente, algunos trabajos abordan el cálculo del ARL mediante simulación. Así, las primeras tablas estimadas para este parámetro (Goldsmith y Whitfield, 1961) se obtuvieron con este método. Posteriormente, Bissell (1969) deriva nomogramas de diseño del gráfico CUSUM con la misma técnica, cuando la distribución de la característica es normal y en presencia de cierta asimetría y curtosis. Van Dobben de Bruyn (1968, p.66) discute la utilización de técnicas de simulación, recomendándolas para el estudio de la distribución del RL cuando el ARL correspondiente es pequeño.

La simulación de Monte Carlo presenta frente a las aproximaciones numéricas antes mencionadas, una mayor flexibilidad y simplicidad, sin pérdida de precisión, siempre y cuando el número de pruebas que se realice sea el apropiado. Además de ésta, otras ventajas son que el programa de ordenador para la simulación es bastante sencillo, no requiere la utilización de ningún lenguaje específico, ya que en todos es posible encontrar una rutina de generación de números aleatorios. Por otro lado, junto con la estimación del ARL (o su intervalo de confianza), este método permite estudiar fácilmente la distribución de RL.

Van Dobben de Bruyn (1968) señala como desventaja para el método de Monte Carlo el tiempo que requiere una simulación, si se desea una precisión buena en el resultado. Obviamente esta observación, por importante que pudiera ser hace veinte años, ha perdido toda relevancia hoy en día por el desarrollo espectacular de la velocidad de cálculo en los modernos ordenadores.

Los métodos que se han mencionado hasta aquí se han utilizado para el estudio de las propiedades y efectividad del gráfico CUSUM en el caso de desviación constante en la media del proceso. Existe tan sólo una referencia en la que se aborde la misma tarea para el caso de desviación progresiva en la media. Bissell (1984a) plantea esta

cuestión, tal y como se mencionó en el apartado de señales adicionales en gráficos de medias, con el fin de comparar distintos procedimientos; la validez de sus resultados es discutible puesto que no se ajustan los ARL bajo control de los gráficos a comparar.

Si con desviación constante es posible abordar el cálculo aproximado del ARL con métodos numéricos, cuando se produce una deriva progresiva de una cierta magnitud en la media el recurso a técnicas de simulación es imprescindible. En este caso no se pueden aplicar las aproximaciones antes comentadas.

Bissell intenta modelizar el gráfico CUSUM según una cadena de Markov no homogénea. Un Corrigendum aparecido en 1986 muestra que la aproximación de Bissell resulta muy poco precisa cuando la deriva es de poca magnitud ( $\leq 0.1$ ). Por ejemplo, para una deriva  $\delta = 0.01\sigma/\sqrt{n}$ , el valor calculado por Bissell para el ARL es igual a 101, y mediante simulación resulta igual a 56.4, con desviación típica 3.8.

En relación con esto último, Bissell (1984b) propone para estimar la magnitud de desviación progresiva a partir del gráfico un método bastante eficiente, pero algo complicado si se desea una gran precisión en la estimación.

Teniendo en cuenta lo anterior, en la presente investigación se ha abordado el cálculo del ARL para el gráfico CUSUM mediante simulación en ordenador, en los casos de desviación constante y de deriva progresiva. El número de pruebas a realizar se discute en el apartado de metodología.

Cuando se produce una señal de falta de control en el gráfico CUSUM, se investiga hasta identificar y eliminar la causa especial de variabilidad, y a continuación se inicializan las sumas acumuladas a cero. Lucas y Crosier (1982) proponen utilizar un valor inicial distinto de cero, ya que puede ocurrir que exista más de una causa especial y que éste no se haya detectado antes de reinicializar el gráfico. El valor inicial permitirá detectar más rápidamente si el proceso está todavía fuera de control. Este nuevo método es denominado FIR CUSUM (Fast Initial Response CUSUM).

Como en cualquier otra variante del gráfico CUSUM, se pretende reducir el ARL a ciertas desviaciones de la media sin reducir significativamente el ARL bajo control, o aumentar la cantidad de falsas alarmas. Lucas et al. muestran que si el valor inicial se elige igual a  $h/2$  se consigue esto último. Por ejemplo, con  $h=5$  el ARL para una desviación igual a  $\sigma/\sqrt{n}$  se reduce de 10.4 a 6.35 en el FIR CUSUM, mientras que el ARL bajo control pasa de 465 a 430. De hecho, el FIR CUSUM con  $h=5$  es incluso más efectivo que el CUSUM con  $h=4$ , pues este último presenta un ARL bajo control igual a 168 y un ARL para desviación 1 igual a 8.38.

Por otro lado, cuando se utiliza el FIR CUSUM, el valor inicial sólo incrementa la efectividad del gráfico si la desviación ya está presente cuando este empieza a funcionar. Si, por otro lado, el proceso está bajo control el efecto de este valor inicial se esfuma rápidamente, funcionando como un CUSUM estándar. En consecuencia, la utilización de esta variante tiene sentido cuando se sospecha que al empezar a operar con el gráfico ya se ha producido una desviación en la media. Por ejemplo, cuando existen múltiples causas de falta de control y sólo se han podido detectar y corregir algunas, o cuando al realizar la corrección de una desviación se ha producido un sobreajuste, etc...

Una ventaja que presenta el gráfico CUSUM respecto del gráfico de medias de Shewhart es la mayor potencia frente a desviaciones de poca magnitud en la media. Ahora bien, el gráfico de medias estándar presenta a su vez una mayor efectividad frente a desviaciones grandes. Con el fin de mejorar en este sentido el CUSUM, Lucas (1982) introduce el procedimiento Shewhart-CUSUM. La diferencia entre este nuevo gráfico y las versiones estándar del CUSUM estriba en que la señal de falta de control viene dada no sólo por las dos sumas acumuladas, sino también por el valor de la observación.

El procedimiento es equivalente a incorporar en el CUSUM límites de control en posiciones  $m_0 \pm z\sigma/\sqrt{n}$ , de forma que se produce una señal de falta de control si la observación sale fuera de estos límites o si alguna de las sumas acumuladas supera el intervalo de decisión  $h$ .



Lucas proporciona algunos resultados para este nuevo gráfico. Por ejemplo, para el CUSUM con  $k=0.5$   $h=5$  (ARL bajo control igual a 465), si se elige  $z=3$ , el gráfico Shewhart-CUSUM resultante presenta un ARL bajo control igual a 223.4, mejorando algo su efectividad frente a desviaciones altas. Por lo tanto, la modificación propuesta se manifiesta, más que en una mejoría notable de potencia a altas desviaciones, en un descenso considerable del ARL bajo control.

Ante esto, Lucas aconseja utilizar valores más altos para  $z$ , como 3.5; el Shewhart-CUSUM con  $k=0.5$ ,  $h=5$  y  $z=3.5$  presenta un ARL bajo control igual a 391, pero en este caso las diferencias con respecto al CUSUM, para desviaciones grandes son incluso más pequeñas que antes. Por otro lado, la introducción de límites de control en el FIR CUSUM no tiene ningún efecto sobre el ARL para desviaciones superiores a 1, disminuyendo considerablemente el ARL bajo control. Es decir, en este caso el efecto de los límites es justo el contrario del que se pretende.

En conclusión, aunque pueden ser útiles en alguna situación concreta, estas variantes del CUSUM suponen algo más de complejidad en su diseño y manejo, sin que se deduzca a partir de los resultados presentados por Lucas et al.(1982) y Lucas (1982) una mejora muy notable en su efectividad. Por tanto, se considerará en los análisis planteados la versión estándar del gráfico CUSUM.

#### **1.2.4- Gráficos de medias móviles**

Cuando no tiene sentido tomar muestras constituidas por varias unidades se recurre a la toma de medidas individuales de la característica de calidad. Esta situación se da frecuentemente, por ejemplo, en la industria química, en la que la mayor parte de los procesos son de tipo continuo, o bien para control de parámetros, como consumos, etc., de los que se dispone de una observación cada cierto tiempo.

En estos casos, se puede recurrir para control on-line de la producción a un gráfico CUSUM o bien a un gráfico de medidas individuales. Este último no es más que un caso particular del gráfico de medias de Shewhart, en el que en lugar de graficar el valor de un estadístico muestral se registran los valores individuales de la característica de calidad.

Como alternativa a este procedimiento, Roberts (1959) introduce los gráficos de medias móviles. Una descripción sobre su construcción se puede encontrar en el Anejo 1, que a su vez se basa en Grant et al. (1980), Wadsworth et al. (1986) y Ryan (1989). Básicamente, el parámetro de diseño a elegir antes de su implantación es el tamaño de agrupamiento  $n$ , según el cual cada punto no es más que la media de los  $n$  últimos valores registrados. Este parámetro determina qué parte de la historia pasada se utiliza, teniendo el mismo peso en cada punto las  $n$  observaciones consideradas.

Aunque se conoce desde hace mucho tiempo e incluso se utiliza de forma habitual en industrias de procesos, resulta cuanto menos sorprendente la inexistencia de análisis sobre el comportamiento y propiedades de este tipo de gráfico en función del tamaño de agrupamiento elegido.

En este procedimiento los valores que se registran están correlacionados entre sí, a diferencia del gráfico de medias estándar, en el que si las observaciones son independientes, ocurre lo mismo con las medias graficadas. Surge por tanto la necesidad del estudio de cuestiones como el tamaño de agrupamiento óptimo o las propiedades estadísticas de estos gráficos, expresadas mediante el valor del ARL correspondiente a distintas situaciones.

Con respecto a este último punto, hay que señalar que para el cálculo del ARL se debe recurrir a técnicas de simulación en ordenador, ya que resulta imposible su determinación analítica exacta. Un artículo reciente de Böhm y Hackl (1990) aborda el cálculo aproximado de algunas cotas superior e inferior del ARL de procedimientos basados en sumas ponderadas de un número finito de valores, siendo el gráfico de medias móviles un caso particular de estos, en el que se pondera igualmente las  $n$

últimas observaciones. Los resultados obtenidos no son útiles para comparar la efectividad de estos procedimientos en función de  $n$ .

Roberts (1959) estudia los ARL del gráfico de medias móviles con  $n=4$ , mediante simulación de Monte Carlo. En otras referencias (Ewan, 1963; Yamane, 1964; Grant et al., 1980) se discute la utilización adecuada de las medias móviles. En general señalan que el gráfico de medias móviles resulta especialmente adecuado para control de características en procesos continuos de la industria química, debido a que la media móvil "suaviza", e incluso puede llegar a eliminar, tendencias o ciclos presentes en los datos.

Por otro lado, Nelson (1983) señala que la representación de medias móviles va ligada a la presencia de pautas y tendencias que pueden dificultar la interpretación de lo que pasa en el proceso. Ilustra esta afirmación representando una serie de 100 valores, independientes y generados a partir de la distribución normal de media cero y desviación típica 1, agrupándolos con tamaño  $n=4$ , 12 y 52.

Teniendo en cuenta lo anterior, una de las cuestiones abordadas en la presente investigación, ha sido, tal y como se explica en el apartado de Objetivos, el estudio de las propiedades estadísticas de los gráficos de medias móviles, en función del tamaño de agrupamiento y la magnitud de desviación en la media. Adicionalmente se compara su efectividad con la de otros procedimientos también utilizados en el contexto de observaciones individuales.

### **1.2.5- Gráficos de control y datos correlacionados**

Los procedimientos para control on-line de procesos considerados en la presente revisión bibliográfica, asumen independencia entre las observaciones a registrar. Ahora bien, esta hipótesis no siempre se cumple. En industrias de procesos, caracterizadas por procedimientos de fabricación continuos o en batch, es frecuente la existencia de correlación entre medidas sucesivas.

En la industria de "partes" (automóvil, electrónica, etc..) si la frecuencia de muestreo es pequeña, también pueden aparecer pautas sistemáticas, debidas a la inercia presente en cualquier proceso de fabricación. Estas pautas o tendencias forman parte de la variabilidad debida a causas comunes.

La falta de independencia en las observaciones puede provocar cambios importantes en las propiedades del gráfico de control elegido, aunque este efecto ha sido muy poco estudiado hasta la fecha.

Los primeros artículos sobre el tema se refieren al gráfico CUSUM. Goldsmith y Whitfield (1961) estudian mediante simulación su comportamiento bajo control, asumiendo para las observaciones  $X_t$  el modelo:

$$X_t = m + \xi_t$$

$$\xi_t = \phi \xi_{t-1} + a_t$$

donde  $m$  es la media del proceso, sujeta a cambios de nivel constantes, y  $a_t$  es la perturbación aleatoria. Los parámetros de la máscara en  $V$  se eligen de forma que resulta un test asimétrico. Concluyen que, si la autocorrelación no es muy grande y el ARL bajo control es pequeño, su valor no se ve muy modificado. Por el contrario, si el ARL es grande, con correlación positiva su valor disminuye, y cuando es negativa aumenta.

Posteriormente, Johnson y Bagshaw (1974 ) y Bagshaw y Johnson (1975) abordan el mismo problema, aproximando el gráfico a un proceso de Wiener y comprobando los resultados mediante simulación. Los modelos que asumen son el autorregresivo de orden uno AR (1) y el de medias móviles, también de orden uno MA(1). Sin embargo, estudian el efecto de la correlación para el procedimiento de sumas acumuladas propuesto inicialmente por Page que, aunque equivalente, no se corresponde con el CUSUM que se utiliza actualmente, ya que supone un valor de referencia  $k=0$ .

Los autores mencionados determinan la modificación del ARL bajo control para un par de valores del coeficiente de correlación (-0.1821 y 0.2195). Concluyen que si el

ARL bajo control es igual a 255.3 con observaciones independientes, cuando la correlación es igual a -0.1821 y 0.2195 pasa a valer respectivamente 360.6 y 187.8. Sin embargo, falta cuantificar cómo se modifica este parámetro frente a desviaciones en la media.

En consecuencia, los estudios referentes a las propiedades y efectividad del gráfico CUSUM deben ser extendidos para un mayor rango de correlaciones y modelos, así como para distintas situaciones de falta de control.

Con respecto al gráfico de medias, Vasilopoulos et al. (1978) estudian el efecto que tiene la correlación en el cálculo de sus límites de control. Proporcionan la expresión correcta para  $\sigma_x$  cuando las observaciones siguen modelos autorregresivos de orden 1 y 2, y tablas con factores para ajustar los límites de control según el valor de la autocorrelación. En general, si se dispone de una serie de medidas individuales que siguen un modelo AR(1) de parámetro  $\phi > 0$ , el valor de la desviación típica de la variable es igual a :

$$\sigma_{X_i} = \frac{\sigma_a}{\sqrt{(1-\phi_1^2)}}$$

donde  $\sigma_a$  la desviación típica de las perturbaciones aleatorias. Si las observaciones  $X_i$  siguen un proceso de medias móviles de orden 1, MA(1), el valor de  $\sigma_x$  resulta ser:

$$\sigma_{X_i} = \sigma_a \sqrt{(1+\theta^2)}$$

Si se toman muestras de  $n$  observaciones consecutivas, con cierta frecuencia, siendo  $\rho$  el coeficiente de correlación entre dos medidas sucesivas, la desviación típica de la media muestral ya no es igual a  $\sigma_x/\sqrt{n}$ . En este caso, se debe calcular como:

$$\sigma_{\bar{x}} = \frac{\sigma_{X_i}}{\sqrt{n}} \sqrt{\left[1 + 2\rho \frac{n-1}{n}\right]}$$

Constable y Yen (1991) proponen utilizar esta expresión para construir los límites de control del gráfico de medias, en el caso de muestras de  $n$  observaciones consecutivas y correlacionadas, según un modelo MA(1). A continuación, comparan mediante simulación el gráfico de medias estándar, y el gráfico de medias con límites

ajustados teniendo en cuenta la expresión anterior. Sin embargo, las comparaciones no se realizan según el valor del ARL, sino que consisten en generar 1000 muestras para cada combinación de estos parámetros, y registrar cuántas resultan en señal de falta de control. Concluyen que en el caso de correlación negativa, resulta conveniente la utilización de la expresión correcta para la desviación típica de la media muestral.

Las conclusiones que se exponen en este trabajo son discutibles puesto que los procedimientos que se comparan no han sido ajustados con el fin de tener el mismo ARL bajo control en todos.

Berthouex, Hunter y Pallesen (1978) aplican el gráfico de medias estándar y el CUSUM a observaciones individuales de una planta depuradora de aguas, que presentan correlación. Encuentran que, en esta situación, estos gráficos proporcionan una gran cantidad de falsas alarmas. Proponen como método alternativo, ajustar la serie de observaciones al modelo ARIMA correspondiente (Box y Jenkins, 1970) y a continuación llevar un gráfico de los residuos del modelo, estimados como la diferencia entre el valor observado y el predicho. Este método, sin embargo, puede dificultar la interpretación de las salidas de control, e incluso de lo que está pasando en cada instante en el proceso, al no tener en cuenta la información potencial contenida en las observaciones iniciales.

En la misma línea, Alwan y Roberts (1988) proponen llevar en paralelo un gráfico de causas especiales y otro para las causas comunes de variabilidad. En el primero se registran los residuos del modelo ARIMA ajustado a los datos, y en el segundo las predicciones, sin incluir límites de control.

Montgomery y Mastrangelo (1991) proponen estimar la predicción aplicando un alisado exponencial, EWMA, a la serie de valores observados. El procedimiento en este caso consiste también en llevar dos gráficos en paralelo. Uno es el gráfico de control de los residuos del modelo EWMA (one-step-ahead EWMA prediction errors). El otro es un gráfico, de la serie original sin límites de control, en el que se incluyen las predicciones EWMA y los límites de confianza correspondientes. El valor óptimo del

parámetro para el EWMA se obtiene minimizando la suma de cuadrados de los errores de predicción.

Los gráficos EWMA fueron introducidos por Roberts (1959) y hasta la fecha han sido bastante utilizados en el campo de la economía, pero muy poco como herramientas para control estadístico de procesos en la industria. Algunos trabajos recientes (Hunter, 1986; Crowder, 1989; Lucas y Saccucci, 1990) abordan el estudio de sus propiedades en este contexto. Sin embargo, en todas estas referencias se asume que las observaciones están incorrelacionadas entre sí.

El estudio de la efectividad de los procedimientos anteriormente expuestos para control con datos correlacionados tiene importancia. Sin embargo, resulta relevante abordar previamente el análisis de las propiedades estadísticas de los gráficos estándar, CUSUM y EWMA, cuando la hipótesis de independencia entre las observaciones deja de cumplirse.

Teniendo en cuenta la escasez de estudios al respecto, en la presente tesis se va a analizar su comportamiento en estas situaciones, con el fin de determinar la magnitud del cambio que sufren sus propiedades en función de la salida de control y de la magnitud de la correlación.

### 1.3- OBJETIVOS

El estudio de las cuestiones abordadas en esta tesis se justifica tanto por la ausencia de investigaciones previas que aporten una respuesta a las mismas, como por su interés práctico para la resolución de problemas en industrias de distintos sectores.

Tal y como se señala en el apartado de revisión bibliográfica, existen muchas referencias en las que se analizan distintos aspectos relacionados con el diseño y propiedades de las técnicas estándar de control on-line de procesos. Sin embargo, algunas cuestiones concretas que han surgido en el marco de nuestra colaboración con diferentes empresas, y que pueden determinar la implantación efectiva de estas técnicas, no han sido abordadas todavía. Con el fin de aportar una respuesta a las mismas, se plantean los análisis que a continuación se exponen.

En primer lugar, en aquellas situaciones en las que el gráfico de medias de Shewhart resulta la herramienta adecuada para control de la característica de calidad, surge a menudo la cuestión de determinar el plan de muestreo óptimo en la etapa de diseño previa a su implantación. Esto implica la necesidad de definir un criterio que permita, con una cierta flexibilidad, fijar el tamaño de muestra y el intervalo de tiempo entre dos muestras consecutivas. La posición de los límites de control se tomará, cuando sea posible, igual a la estándar,  $\pm 3\sigma/\sqrt{n}$ . Se facilita así el manejo del gráfico, por la gran cantidad de tablas existentes en manuales y libros para las constantes a utilizar en el cálculo de sus límites.

Por otro lado, y debido a la amplia gama de situaciones distintas en las que este gráfico puede ser aplicado, el estudio de un criterio general, que pueda ayudar a diseñar el plan de muestreo óptimo, resulta en principio una tarea inabordable. En consecuencia, se va asumir alguna simplificación que permita delimitar el problema. La elaboración de aproximaciones más o menos generales ya ha sido analizada por otros autores, resultando en modelos y algoritmos excesivamente complicados, y por lo tanto sólo útiles también en un determinado número de casos particulares.



La elección del plan de muestreo adecuado en un gráfico de control debe basarse en consideraciones tanto económicas como estadísticas. Las primeras permiten cuantificar el esfuerzo total de muestreo en función de parámetros como su coste unitario, la probabilidad de salidas de control y el coste de las unidades defectuosas producidas antes de que se detecten.

En situaciones en las que el coste total de 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 de tiempo frente a otros basados en muestras de mayor tamaño, pero más espaciadas. En estos casos, el parámetro que puede determinar la elección en uno u otro sentido es el tiempo que tarda en detectarse una salida de control, que se manifieste en la práctica como una desviación de magnitud constante en la media del proceso. Así, para una desviación dada, es preferible aquel plan de muestreo que permite en promedio minimizar el intervalo entre dos señales de falta de control, ya que para procesos con una producción horaria constante este criterio reduce el coste debido a unidades defectuosas.

Por lo tanto, la primera cuestión que se ha abordado en la presente tesis ha sido el estudio comparativo de diferentes planes de muestreo, con el fin de analizar cómo dicha elección afecta a la potencia del gráfico de medias, definida ésta como su capacidad para detectar rápidamente salidas de control del proceso. Se considera el procedimiento de medias estándar, en el que la señal de falta de control consiste en la aparición de un punto fuera de los límites. Se ha tomado como referencia para las comparaciones un gráfico con límites a  $\pm 3\sigma/\sqrt{n}$  y tamaño de muestra  $n=5$ . Se asume también que las observaciones son independientes entre sí y que las medias muestrales siguen una distribución aproximadamente normal. Las salidas de control vienen dadas por desviaciones de magnitud constante, que se manifiestan mediante un cambio de nivel en la media del proceso.

Una vez realizada la comparación según lo indicado en el párrafo anterior, se analiza la elección del plan de muestreo óptimo en función de la capacidad del proceso, en los casos de especificación bilateral y unilateral para la característica de calidad.

En la segunda parte de la presente investigación se ha abordado el estudio de otra cuestión también relacionada con la utilización eficiente de los gráficos de medias. En su versión estándar, este procedimiento se caracteriza por una efectividad relativamente baja frente a desviaciones de poca magnitud en la media del proceso. Con el fin de aumentar su potencia en estas situaciones, se introduce en la práctica habitual señales adicionales de falta de control, en las que se considera la información contenida en un cierto número de muestras, y no sólo el último punto graficado, como ocurre en la propuesta inicial de Shewhart.

Tal y como se indicó en la revisión bibliográfica, existen distintas combinaciones de señales adicionales, según la pauta o racha a tener en cuenta. En el presente trabajo, se va a considerar el conjunto de señales adicionales propuesto por Ishikawa (1976), y utilizado de forma más frecuente en algunas industrias españolas (Ford, 1984). Se incluye en el apartado anterior una definición de las mismas.

Pese a que estas señales adicionales aparecen en manuales sobre control de procesos desde hace algún tiempo, resulta sorprendente la carencia de estudios que cuantifiquen exactamente el efecto de las mismas sobre las propiedades estadísticas del gráfico de medias. Una primera consecuencia de su utilización es, evidentemente, una mayor complejidad de manejo del gráfico de medias. Por otro lado, intuitivamente se ve que por el hecho de considerar en estas señales parte de la información pasada aumentará su potencia frente a desviaciones de poca magnitud, pero al mismo tiempo se modificará también la cantidad de falsas alarmas.

Tal vez la carencia de estudios en los que se determine exactamente esto último, se deba a la dificultad de cálculo que conllevan, ya que si entre las señales adicionales se incluye la señal de racha ascendente o descendente, la única forma de abordar el problema es mediante simulación en ordenador.

Teniendo en cuenta lo anterior, y lo expuesto en la revisión bibliográfica, se ha realizado el estudio de las propiedades estadísticas del gráfico de medias cuando se incluyen estas señales adicionales mediante el recurso a técnicas de simulación. Se

consideran distintos tamaños de muestra y desviaciones en la media, tanto de magnitud constante como progresiva.

Uno de los objetivos de este análisis es determinar hasta qué punto estas señales mejoran la potencia del gráfico, teniendo en cuenta que su inclusión supone también una mayor complejidad de manejo. Como medida de la efectividad se utiliza el número promedio de muestras a tomar desde que se produce la salida de control hasta que se manifiesta alguna de las señales, o ARL, que también cuantifica el número de falsas alarmas cuando las señales se producen estando el proceso bajo control.

Con el fin de determinar la variación en potencia que estas señales suponen, se compara según el valor del ARL, el gráfico de medias con señales adicionales con el gráfico de medias estándar y el gráfico CUSUM, en los casos de desviación progresiva y constante en la media del proceso. Previamente, se ajustan los tres gráficos para un mismo ARL bajo control, tomando para el gráfico con señales adicionales los límites a  $\pm 3\sigma/\sqrt{n}$  de la media del proceso.

Adicionalmente, se estudia la frecuencia relativa de aparición de cada señal, para determinar la contribución de cada una en el cambio de propiedades estadísticas en los casos estudiados. Con el fin de analizar la interpretación de estas señales, se estima también esta frecuencia relativa cuando la salida de control se debe a una variación en la dispersión de la característica de calidad a controlar y la media se mantiene constante.

Cuando no tiene sentido plantearse la toma de muestras de un determinado tamaño y se dispone de una sola observación por intervalo de tiempo, se utilizan frecuentemente los gráficos de medidas individuales, que son un caso particular de los gráficos de medias de Shewhart para  $n=1$ .

Como alternativa a este último, y considerando que la dispersión de una media es siempre menor que la de las observaciones individuales, se utilizan también los gráficos de medias móviles. Situaciones de este tipo se dan, por ejemplo, en control de procesos continuos en la industria química, o para control de variables como consumos,

temperaturas, materia prima, etc...

Si se opta por un gráfico de medias móviles se debe elegir, en la etapa previa a su implantación, el tamaño de agrupamiento  $n$  con el que se calculan las medias a graficar. Se debe tener en cuenta que sea cual sea el valor de  $n$ , los gráficos de medias móviles resultantes son equivalentes en cuanto a coste de muestreo, pero difieren en sus propiedades estadísticas.

Tal y como se indica en la revisión bibliográfica, no existe ningún estudio en el que se aborde la determinación de las propiedades estadísticas del gráfico de medias móviles, en función del tamaño de agrupamiento elegido. Por lo tanto, tampoco existe ningún criterio estadístico que permita elegir el  $n$  apropiado en cada situación.

En consecuencia, nos hemos planteado como tercera cuestión a analizar, el estudio de las propiedades estadísticas del gráfico de medias móviles en función del tamaño de agrupamiento y de la magnitud de desviación constante en la media del proceso. Se asume, para la característica de calidad, una distribución aproximadamente normal y que las observaciones son independientes entre sí. La posición de los límites de control se fija inicialmente a  $\pm 3\sigma/\sqrt{n}$ . El parámetro a utilizar como medida de su potencia es también el ARL, que se obtiene mediante simulación por no existir en este caso aproximación analítica para su cálculo.

Como alternativa a los gráficos de medias móviles, se dispone también de los gráficos CUSUM y EWMA, recomendados frecuentemente para control de procesos en la industria química. Teniendo esto en cuenta se compara el gráfico de medias móviles con los dos anteriores y el de medidas individuales, con el fin de determinar el procedimiento óptimo en función de la desviación en la media.

En los análisis anteriormente mencionados, se asume de partida que las observaciones son independientes entre sí. Sin embargo, es frecuente encontrar situaciones en las que esta hipótesis deja de cumplirse por la misma naturaleza del proceso a controlar. Tal es el caso, por ejemplo, de procesos continuos, en los que el mismo

sistema de medida puede provocar una estructura de correlaciones entre observaciones sucesivas.

Según se refleja en la revisión bibliográfica, el control on-line de procesos en los que la información se presenta en una serie de valores correlacionados ha cobrado interés recientemente. Habitualmente se utilizan los gráficos estándar, pese a que en su fundamento asumen que las observaciones son independientes entre sí.

Resulta, por tanto, relevante determinar el comportamiento estadístico de los gráficos más utilizados en la práctica cuando los datos a registrar presentan una estructura de este tipo.

En la bibliografía sobre el tema, no hemos encontrado respuesta a este problema, más que en un número muy reducido de casos particulares, que se indican en el apartado anterior. Por tanto, la última cuestión planteada en la presente investigación consiste en estudiar las propiedades estadísticas del gráfico de medidas individuales, de medias móviles, CUSUM y EWMA, cuando se aplican a observaciones correlacionadas.

Se considerará su comportamiento cuando el proceso está bajo control, con el fin de cuantificar cómo se modifica la cantidad de falsas alarmas y cuando tiene lugar una desviación de magnitud constante en la media. Se pretende comparar la efectividad y propiedades de cada gráfico con las correspondientes a su aplicación a observaciones independientes.

## **CAPÍTULO 2**

### **METODOLOGÍA**

#### **2.1- COMPARACIÓN DE GRÁFICOS DE CONTROL: ARL**

En las cuestiones consideradas en esta investigación, se pretende evaluar las características operativas de gráficos de control para la media bajo distintas hipótesis referentes a salidas de control y estructura de las observaciones. Se precisa por tanto de algún parámetro que permita cuantificar la efectividad y potencia de estos procedimientos.

Cuando se considera el gráfico de medias estándar, por su equivalencia con un contraste de hipótesis sobre la media, resultan relevantes los conceptos de riesgo de primera y segunda especie. El primero es la probabilidad de concluir que la media ha sufrido una desviación cuando en realidad sigue bajo control, siendo equivalente a la probabilidad de falsa alarma o punto fuera de los límites de control en estas circunstancias.

Debido a que en este gráfico se asume independencia entre las observaciones es posible calcular el valor de  $\beta$  fácilmente. Cuando se produce una desviación en la media la potencia del gráfico viene dada por  $(1-\beta)$ , siendo  $\beta$  el riesgo de segunda especie o probabilidad de que no se produzca señal de falta de control.

En relación con lo anterior, se considera la longitud de recorrido (run length), o número de muestras a tomar desde que se produce la salida de control en el proceso hasta que tiene lugar la señal en el gráfico.

Cuando el proceso está bajo control, la longitud de recorrido cuantifica el número de muestras tomadas entre dos falsas alarmas. Se trata, por tanto, de una

variable aleatoria, discreta y que en el caso del gráfico de medias estándar sigue una distribución geométrica.

En efecto, si  $p$  es la probabilidad de que un punto caiga fuera de los límites de control, se tiene:

$$P(RL=r) = (1-p)^{r-1} p$$

El valor esperado de esta variable, también conocido como longitud de recorrido medio (Average Run Length) o ARL, es igual a  $1/p$ . La desviación típica resulta ser igual a:

$$\sigma_{RL} = \frac{\sqrt{(1-p)}}{p}$$

y teniendo en cuenta lo anterior

$$\sigma_{RL} = \sqrt{(1-p)} ARL$$

Si el valor de  $p$  es pequeño, situación deseable cuando el proceso está bajo control, se puede asumir:

$$\sigma_{RL} \approx ARL$$

por lo que la desviación típica de la longitud de recorrido es aproximadamente del mismo orden de magnitud que su media.

La relación de este parámetro con los riesgos de primera y segunda especie es directa, puesto que si el proceso está bajo control

$$ARL = 1/\alpha,$$

y si está fuera de control

$$ARL = 1/(1-\beta)$$

En consecuencia, tal y como se reflejó en la revisión bibliográfica, tiene sentido utilizar el valor del ARL como medida de la efectividad del gráfico de medias estándar y para realizar comparaciones en ciertos contextos.

El valor de p se calcula en este gráfico como:

$$p = P(|\bar{X}_i - m_0| > 3\sigma/\sqrt{n})$$

Si se asume, como es lo habitual, distribución aproximadamente normal para la característica de calidad, o más concretamente para las medias muestrales, la expresión anterior se evalúa según esta distribución de probabilidad. En el presente trabajo se ha utilizado para cálculo de probabilidades de esta función de distribución de probabilidad acumulada, la subrutina proporcionada por Craig (1984).

El gráfico de medias con límites a  $\pm 3\sigma/\sqrt{n}$  implica un  $\alpha = 0.0027$  y por lo tanto un ARL bajo control igual a 370. Esto quiere decir que en promedio de cada 370 observaciones 1 resulta en una falsa alarma. Dicho de otra forma, en promedio habrá 370 observaciones entre dos falsas alarmas consecutivas.

El ARL también se denomina ASRL (Average Sample Run Length) para distinguirlo del AARL, o número promedio de artículos muestreados desde que se produce la salida de control hasta que se detecta en el gráfico. Evidentemente, si el tamaño de muestra es constante, que es lo habitual en la práctica, e igual a n, AARL = n x ARL. Este parámetro puede resultar más apropiado que el ARL cuando se comparan procedimientos que difieren en el valor de n y, por tanto, pueden presentar distintos costes de muestreo.

Si además del tamaño de muestra, varía también el intervalo de tiempo entre dos muestras consecutivas resulta relevante en la comparación de procedimientos considerar el número total de artículos defectuosos producidos. Esto último es equivalente, cuando la producción horaria es constante, al tiempo que se tarda en detectar la salida de control.

En relación con esto, se define como medida de la efectividad del gráfico el AvD (Average Duration), que no es más que el valor promedio de ese tiempo. Si f es la frecuencia de muestreo, o intervalo de tiempo entre dos muestras consecutivas, y t es el tiempo promedio entre el instante en el que se produce la desviación y la toma de la



siguiente muestra, el AvD viene dado por:

$$AvD = (ARL - 1) \times f + t$$

Si la probabilidad de que el cambio en el proceso tenga lugar en algún intervalo de tiempo concreto es pequeña, puede asumirse que la salida de control se produce, en promedio, en el punto central del intervalo entre dos muestras consecutivas, con lo que  $t \approx f/2$ , resultando

$$AvD = (ARL - 0.5) \times f \quad (1)$$

Este parámetro es relevante para la primera cuestión a abordar en nuestra investigación. Cuando se puede asumir que el coste total de 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 menos frecuentes.

En estos casos, el plan de muestreo óptimo es aquel que minimiza en promedio el intervalo de tiempo que tarda en detectarse una salida de control real en el proceso, ya que para un mismo coste de muestreo, si la producción horaria es constante las pérdidas debidas a unidades defectuosas serán también mínimas.

Por lo tanto, para la comparación entre planes de control en este caso se utilizará el valor del AvD.

Puesto que se comparan planes que difieren tanto en el tamaño de muestra  $n$  como en la frecuencia de muestreo  $f$ , la magnitud de desviación constante de la media del proceso viene expresada en unidades de desviación típica del mismo, que se asume conocida y constante.

Con el fin de cubrir la gama de situaciones más frecuentes en la práctica, se considera el procedimiento consistente en tomar observaciones individuales muy frecuentemente, definiendo en este caso  $f=1$  Ud tiempo arbitraria, y el plan consistente en tomar muestras de tamaño relativamente grande  $n=7$ , pero con muy poca frecuencia,

por lo que se toma el intervalo de muestreo  $f=7$ . Como planes intermedios se incluyen  $n=3$  con  $f=3$  y  $n=5$  con  $f=5$ . Teniendo en cuenta la simplificación asumida, todos los procedimientos resultan equivalentes desde el punto de vista de coste.

Adicionalmente se han ajustado los límites de control de forma que el tiempo promedio (AvD) que tarda en producirse una falsa alarma sea el mismo en todos los casos. En concreto, se ha tomado como referencia el AvD bajo control del gráfico de medias con tamaño de muestra  $n=5$ ,  $f=5$  y límites de control en la posición estándar  $m_0 \pm 3\sigma/\sqrt{n}$ .

Una vez ajustados los planes a comparar, el procedimiento de cálculo ha consistido básicamente en suponer una determinada desviación en la media, y calcular el ARL correspondiente. A partir de este, aplicando la expresión (1) se determina el valor de AvD para cada caso. Con los resultados de esta primera parte, se analiza el tamaño de muestra óptimo en función de la capacidad del proceso y de la magnitud de desviación cuya detección precoz se considera relevante, dependiendo esta última del porcentaje de unidades defectuosas que resultan de la salida de control.

Como segundo objetivo de la presente tesis, se pretende estudiar las propiedades estadísticas del gráfico de medias con un determinado conjunto de señales adicionales de falta de control. Estas señales se incluyen con el fin de mejorar la potencia del gráfico frente a desviaciones de poca magnitud. En este caso se utiliza también el ARL como medida de la efectividad, aunque tal y como se indica en el apartado de revisión bibliográfica, su determinación analítica resulta mucho más complicada.

En primer lugar, los conceptos de riesgo de primera y segunda especie dejan de tener sentido. La probabilidad de señal de falta de control depende del instante de tiempo en que nos encontremos, ya que las señales se basan en parte de la información contenida en observaciones anteriores. Adicionalmente, estas señales no son independientes entre sí.

Las aproximaciones propuestas para el cálculo analítico del ARL de este procedimiento resultan extremadamente complejas, e incluso son inabordables cuando se incluye la señal ascendente o descendente de siete puntos. En consecuencia, se ha recurrido a la simulación en ordenador para la determinación del ARL en las distintas situaciones de falta de control.

Con los otros gráficos a estudiar (CUSUM, medias móviles, etc..) también se ha utilizado la simulación para el cálculo de sus propiedades, expresadas en función del ARL bajo control y cuando se produce una desviación en la media. En el siguiente apartado se expone el procedimiento seguido para los distintos gráficos de control.

## **2.2- PROCEDIMIENTOS PARA CÁLCULO DEL ARL**

Exceptuando el gráfico de medias estándar, no existe ningún procedimiento analítico que permita determinar exactamente el valor del ARL de los otros gráficos más utilizados para control de la media.

Los métodos aproximados para el caso del gráfico de medias con señales adicionales resultan extremadamente complicados cuando aumenta el número de señales, siendo inabordables cuando se considera la señal de racha ascendente o descendente. Por tanto, se determinará el valor del ARL mediante el recurso a técnicas de simulación.

Con respecto al gráfico CUSUM, cuando la magnitud de desviación es constante es posible evaluar sus propiedades expresadas como el ARL bajo control o fuera de control, según la situación, resolviendo de forma aproximada las ecuaciones integrales correspondientes a este parámetro, o bien modelizando el gráfico según una cadena de Markov.

Una tercera posibilidad es calcular el ARL mediante simulación. Este método presenta frente a las aproximaciones numéricas antes mencionadas una mayor flexibilidad y simplicidad, sin pérdida de precisión, tal y como se expone más adelante en la discusión sobre el número de pruebas a realizar.

Además de éstas, otra ventaja del enfoque de simulación es que el programa de ordenador preciso es bastante sencillo y no requiere ningún lenguaje en particular, ya que todos disponen de una rutina para generación de números aleatorios. Es posible también determinar la precisión de la estimación dada para el ARL en cada caso, mediante el cálculo de su desviación típica. Por otro lado, se puede estudiar fácilmente la distribución de la variable longitud de recorrido, tarea que para el gráfico CUSUM resulta inabordable de forma analítica. En el caso de desviación progresiva en la media del proceso el cálculo del ARL de todos los procedimientos ha sido abordado también mediante simulación, dado que este es el único procedimiento posible.

En la tercera cuestión analizada en esta tesis se estudian las propiedades estadísticas de los gráficos de medias móviles en función del tamaño de agrupamiento y de la magnitud de desviación constante de la media del proceso. En este tipo de gráfico las observaciones que se registran no son independientes entre sí.

Se demuestra fácilmente que el coeficiente de correlación entre dos medias  $i$  e  $i-j$  es igual a  $(n-j)/n$ , para  $j < n$ , siendo  $n$  el tamaño de agrupamiento elegido. Por tanto, no es posible aplicar la expresión vista para el cálculo del ARL en el caso del gráfico de medias estándar. El cálculo analítico de este parámetro resulta también en este caso extremadamente complicado, tal y como se señala en la revisión bibliográfica. Teniendo en cuenta todo lo anterior, se va a proceder a estimar las propiedades estadísticas mediante simulación del gráfico, según el procedimiento que se expone más adelante.

Finalmente, para el estudio del comportamiento estadístico de los gráficos aplicados a observaciones correlacionadas entre sí, se ha recurrido también a la simulación en ordenador de distintas situaciones en las que varían la magnitud de la desviación en la media y el valor del coeficiente de correlación de los valores a registrar. En cada caso, la serie de observaciones ha sido generada mediante simulación asumiendo un determinado modelo.

Si bien para el caso de observaciones independientes existe alguna aproximación al cálculo analítico del ARL para algún gráfico concreto, en presencia de correlación todavía no ha sido abordada esta cuestión. El objetivo de esta parte de nuestra investigación no es el estudio de métodos de cálculo del ARL, sino la determinación de valores del mismo de la forma más eficiente posible, con el fin de cuantificar la modificación de las propiedades de los gráficos. Por tanto, también se ha abordado esta cuestión mediante simulación en ordenador.

### 2.3- GENERACIÓN DE LAS OBSERVACIONES

La distribución de probabilidad asumida para la característica de calidad en todos los casos es la normal. Para algunas de las cuestiones anteriormente mencionadas, se requiere de un procedimiento que permita obtener observaciones que sigan este modelo y sean independientes entre sí.

Si la media del proceso es igual a  $m$  en un instante  $i$  cualquiera y la desviación típica es  $\sigma$ , una observación individual puede expresarse en función de esta media y de la perturbación aleatoria correspondiente de la forma:

$$X_i = m + \varepsilon_i \sigma$$

donde

$$\varepsilon_i \sim N(0, 1)$$

y que asumimos independientes entre sí.

Para la última cuestión a abordar se precisa también de algún procedimiento para generar observaciones correlacionadas. Si, por ejemplo, se asume que las observaciones independientes siguen un modelo autorregresivo de orden uno, AR(1), se tiene que en el instante  $i$ , la observación  $X_i$  puede expresarse como:

$$X_i - m = \phi(X_{i-1} - m) + a_i$$

donde  $a_i$  son las perturbaciones aleatorias, con distribución normal  $N(0, \sigma_a)$ . Sin pérdida de generalidad, asumimos que  $\sigma_a$  es igual a 1, y que el valor inicial de la serie  $X_0 = a_0$ .

Por tanto, como paso previo a la realización de cualquiera de estas simulaciones, se ha estudiado la manera de generar una serie de perturbaciones aleatorias con distribución  $N(0,1)$  e independientes entre sí.

El procedimiento seguido consiste en aplicar la transformación de Box y Muller a valores de una distribución uniforme  $U(0,1)$ . Así, si  $U_1$  y  $U_2$  son dos variables aleatorias, independientes que siguen esta última distribución, se demuestra fácilmente que las expresiones:

$$X = \sqrt{-2 \ln U_1} \cos 2\pi U_2$$

$$Y = \sqrt{-2 \ln U_1} \sin 2\pi U_2$$

proporcionan dos valores  $X$  e  $Y$  independientes con distribución normal tipificada.

El lenguaje empleado en las simulaciones es el APL, versión 2 de IBM, que entre otras cosas no presenta limitación de tamaño de objeto y permite gestionar de forma eficiente matrices y vectores de gran longitud. Para la generación de los valores  $U(0,1)$  independientes, se utiliza la función "azar" de este lenguaje, que aplicada a un número cualquiera  $n$ , da como resultado un valor al azar entre 0 y  $n$ . El programa utilizado para generar una serie de perturbaciones aleatorias de longitud determinada se recoge en el Anejo 2.

## **2.4- ANÁLISIS DEL VECTOR BÁSICO DE PERTURBACIONES**

Mediante el procedimiento antes descrito, se generó un vector de perturbaciones aleatorias de longitud 10000, y se estudiaron sus propiedades, con el fin de analizar la calidad del generador de números pseudoaleatorios.

En primer lugar se comprobó que no existe ciclicidad en este vector. Esta propiedad es importante puesto que si el número de pruebas en cada simulación es grande, se requiere que la longitud de la secuencia de números generados consecutivamente al azar sea larga.

Con el tipo de ordenador utilizado en las simulaciones, el periodo máximo que se puede obtener con un método congruencial de generación de métodos pseudoaleatorios es igual a  $2^{32}$ , valor que asegura una longitud de ciclo suficientemente larga para las simulaciones a realizar en nuestro trabajo. Por otro lado se comprobó la no existencia de ciclicidad, aplicando el programa que se recoge en el Anejo 2 al vector básico de perturbaciones.

Adicionalmente, se han estudiado las características de este vector de 10000 valores, con el fin de comprobar que su comportamiento es similar al de un vector de valores  $N(0,1)$  independientes.

A partir de este vector, se obtuvo una muestra al azar de 1000 valores que fueron analizados con el STATGRAPHICS. La FIGURA 2.1 recoge una representación gráfica de las frecuencias de valores observados en esta muestra y la TABLA 2.1 sus parámetros descriptivos.

Se ha realizado también un test Chi-Cuadrado sobre la bondad del ajuste a una distribución  $N(0,1)$  de los valores contenidos en el vector, tomando 14 intervalos. El estadístico correspondiente resulta igual a 11.825, con un nivel de significación 0.3769, por lo que es admisible que el vector de 1000 valores analizado procede de una población con distribución  $N(0,1)$ .

La independencia entre las perturbaciones generadas se comprueba con el correlograma del vector, tal y como se refleja en la FIGURA 2.2, en el que no aparece ningún coeficiente de correlación simple significativo.

Una vez obtenidas las perturbaciones aleatorias, y hechas las comprobaciones anteriores, se puede proceder a generar las observaciones correspondientes a cada caso, según lo explicado antes. El procedimiento general a seguir ha consistido en simular una salida de control y registrar el número de muestras generadas hasta que se produce la señal de falta de control en el gráfico correspondiente.

FIGURA 2.1- Histograma de frecuencias para el vector de perturbaciones de longitud 10000, generado según el método de Box y Muller.

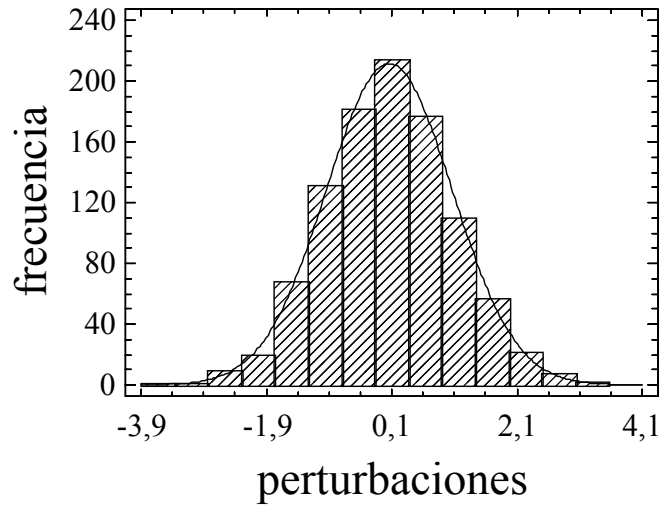
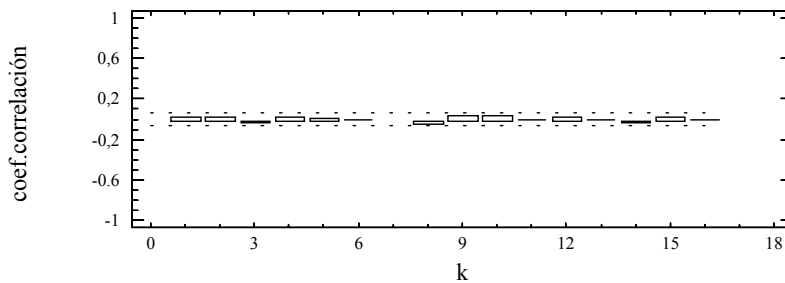


FIGURA 2.2- Correlograma del vector de 1000 valores  $N(0,1)$



## 2.5- DETERMINACIÓN DEL NÚMERO DE REPETICIONES

Puesto que se pretende estimar a partir de estos resultados la media de la variable longitud de recorrido, tiene importancia fijar el número de repeticiones a ejecutar en cada situación particular, con el fin de estimar este parámetro con precisión suficiente. La longitud de recorrido sigue una distribución geométrica de parámetro  $1/ARL$ , cuando el procedimiento de control es el gráfico de medias estándar. En cualquier otro caso, v.g. gráfico CUSUM, su distribución es claramente asimétrica por la



derecha y se puede aproximar sin mucho margen de error a la geométrica cuando el ARL correspondiente es grande.

TABLA 2.1- Estadísticos muestrales de los 1000 valores extraídos al azar del vector de perturbaciones aleatorias.

Media	0.0612
Mediana	0.0690
Moda	0.1893
Desviación típica	1.0048
Primer Cuartil	-0.6291
Tercer Cuartil	0.7018
Intervalo intercuartílico	1.3309
Asimetría	0.0223
Asimetría estandarizada	0.2877
Curtosis	0.0774
Curtosis estandarizada	0.4995
Mínimo	-3.5552
Máximo	3.1923

Como se vio en el apartado 2.1, la desviación típica de la variable longitud del recorrido es aproximadamente igual al ARL, cuando el valor de éste es grande. En cualquier otra situación será menor. Teniendo esto en cuenta, vamos a asumir  $\sigma_{\text{ARL}} = \text{ARL}/\sqrt{N}$ , por ser el caso más desfavorable, con el fin de fijar el tamaño N de simulación adecuado.

Se pretende estimar un valor ARL, a partir de N simulaciones de un gráfico de control cualquiera, registrando en cada prueba el número de muestras a tomar hasta que se produce señal de falta de control en el mismo. Si se desea que la media de estos N valores, que será el ARL muestral, no difiera del verdadero valor en más de un 2% del mismo con una probabilidad del 95% se tiene:

$$P(|\overline{ARL} - ARL| < 0.02 ARL) \geq 0.95$$

y suponiendo que N es lo bastante grande para poder asumir distribución normal para el ARL muestral, resulta que  $N \geq 9604$ .

Redondeando el número de simulaciones a  $N=10000$ , se tendrá que con una probabilidad del 95%, el ARL estimado no diferirá del verdadero valor en más de 1.96% del mismo. Análogamente, con una probabilidad del 99%, la diferencia entre la estimación y el verdadero valor no supera el 2.5% de este último. Por otro lado, teniendo en cuenta que la desviación típica de la estimación es aproximadamente del orden de  $ARL/\sqrt{N}$ , su coeficiente de variación no superará en ningún caso el 1%.

Una vez elegido el número de pruebas, se ha procedido a simular cada procedimiento según lo que se explicita a continuación.

Con el fin de ilustrar la precisión de los resultados obtenidos, se recoge en la TABLA 2.2 valores del ARL del gráfico de medias estándar con tamaño de muestra 5, determinados exactamente y mediante simulación con un número de pruebas igual a 10000.

Según se aprecia en la tabla mencionada, las estimaciones para el ARL mediante simulación no difieren significativamente de los valores exactos. Las diferencias más grandes aparecen para desviaciones en la media de poca magnitud, aunque esto tiene sentido, puesto que en este caso la desviación típica de la variable longitud de recorrido es también mayor.

En la TABLA 2.3 se recoge una comparación entre los ARL bajo control del gráfico CUSUM estimados analíticamente, según los métodos mencionados en la revisión bibliográfica, y los ARL bajo control obtenidos con el procedimiento de simulación que más adelante se explicita, con un número de pruebas igual a 10000.

TABLA 2.2- Comparación de valores del ARL para el gráfico de medias estándar con tamaño de muestra 5. Entre paréntesis la desviación típica de la estimación.

$\delta$ (Ud $\sigma$ )	ARL Exacto	ARL Simulación
0.00	370.39	364.88 (3.62)
0.25	133.16	135.09 (1.36)
0.50	33.40	33.39 (0.33)
0.75	10.76	10.59 (0.11)
1.00	4.49	4.47 (0.04)
1.25	2.38	2.39 (0.03)
1.50	1.56	1.57 (0.01)
1.75	1.22	1.23 (0.01)
2.00	1.07	1.07 (0.01)
2.25	1.02	1.02 (0.01)
2.50	1.00	1.00 (0.01)

Se observa también en este caso que los valores estimados a partir de la simulación no presentan diferencias significativas con los obtenidos por otros autores mediante aproximación numérica. Por lo tanto, la precisión conseguida con el número de repeticiones elegido se considera suficientemente buena.

Con respecto al tiempo que requiere cada simulación, para estimar un ARL del orden de 370, se precisan con un ordenador IBM PS/2 serie 80, con el programa escrito en APL, unas cinco horas de cálculo, dependiendo esto del gráfico simulado. Por supuesto, esta es la situación más desfavorable, puesto que los ARL a estimar en los análisis abordados son en la mayor parte de los casos muy inferiores, y la duración de cada simulación particular es proporcional al RL de la misma.

TABLA 2.3- ARL bajo control para el gráfico CUSUM, estimado según distintos métodos.

CUSUM k=0.5 h=3			
Brook y Evans (1972)	Woodall (1983)	Waldmann (1986)	Simulación
58.6	58.8	58.7	59.05 (0.56)
CUSUM k=1 h=2.5			
Goel y Wu (1972)	Lucas (1976)	Woodall y Adams (1991)	Simulación
358	357.9	352.46	358.07 (3.55)

## 2.6- SIMULACIÓN DEL FUNCIONAMIENTO DE LOS DISTINTOS GRÁFICOS

### 2.6.1- Simulación de la salida de control

Los programas en APL para la simulación de cada gráfico de control se recogen en el Anejo 2. La magnitud de desviación se asume constante o progresiva. Cuando la desviación  $\sigma$  es constante, la media del proceso se mantiene igual a  $m_0 + \delta\sigma$  hasta que se detecta. El valor  $m_0$  es el nivel inicial antes del cambio.

Si es progresiva, se considera que la tendencia es lineal, por lo que la media de las observaciones tomadas en la  $i$ -ava muestra después de producirse la desviación es igual a  $m_0 + \delta i \sigma$ . En este caso, la magnitud de desviación  $\delta \sigma$  expresa la tasa de desviación de la media por intervalo de muestreo. La simulación de una salida de control debida a una desviación progresiva supone sólo una pequeña modificación del programa correspondiente a desviación constante en la media.

Se programarán los procedimientos con observaciones estandarizadas, de forma que si  $X_{ij}$  es la  $j$ -ava medida de la muestra  $i$ , se parte del valor:

$$\frac{X_{ij} - m}{\sigma} = \varepsilon_{ij}$$

que no es más que la perturbación aleatoria asociada a ese momento. Por tanto, para generar muestras de tamaño  $n$ , se tomarán  $n$  valores al azar del vector de residuos  $N(0,1)$  de longitud 10000. Si la muestra corresponde a un instante de tiempo en el que el proceso está bajo control, el promedio de estas  $n$  perturbaciones será la media a contrastar con las señales de falta de control. Si por el contrario se ha producido una desviación de magnitud  $\delta \sigma$ , a la media de los  $n$  residuos se le suma la cantidad  $\delta$  si la desviación es constante. En la hipótesis de desviación progresiva se suma la cantidad  $i$ , siendo  $i$  el número de muestra desde que tuvo lugar la salida de control.

## 2.6.2- Simulación del gráfico de medias

Para el caso concreto de simulación del gráfico de medias con señales adicionales, el trabajar con observaciones estandarizadas implica una línea central situada en cero, límites de control a  $0 \pm 3/\sqrt{n}$ , y límites de aviso a  $0 \pm 2/\sqrt{n}$ .

Las 4 señales de falta de control se transforman en:

**Señal 1:** Media muestral supera en valor absoluto el valor  $3/\sqrt{n}$ .

**Señal 2:** Dos de las tres últimas medias están en valor absoluto por encima de  $2/\sqrt{n}$ .

**Señal 3:** Las siete últimas medias tienen el mismo signo.

**Señal 4:** Las seis diferencias entre las siete últimas medias son del mismo signo.

Teniendo en cuenta lo anterior, el diagrama de flujo del programa es el que se recoge en la FIGURA 2.3. Tal y como se observa en la figura mencionada, el resultado de la simulación es un vector de 10000 observaciones de la variable longitud de recorrido, a partir del cual se puede determinar el ARL como el valor medio de las mismas, su desviación típica y otros parámetros relacionados con la distribución de esta característica. Para estudiar el ARL bajo control, se da como input para la desviación el valor  $\delta=0$ . Adicionalmente, se registra también el número de veces que se presenta cada señal de falta de control, por lo que es posible estudiar la frecuencia relativa de cada una en función de la desviación o del tamaño de muestra.

### **2.6.3- Simulación del gráfico de medias móviles**

Para la simulación del gráfico de medias móviles también se ha trabajado con observaciones estandarizadas. En concreto, la media  $i$  se obtiene promediando la observación individual del instante  $i$  con  $n-1$  observaciones anteriores, siendo  $n$  el tamaño de agrupamiento. Por tratarse de valores estandarizados, cada observación no será más que un valor obtenido al azar del vector de perturbaciones aleatorias. La línea central del gráfico se sitúa en cero y los límites de control a  $\pm 3/\sqrt{n}$ .

La FIGURA 2.4 recoge el diagrama de flujo del programa de simulación de este gráfico.

Como más adelante se indica, también se ha cuantificado la aparición de rachas en este procedimiento. El programa en este caso en lugar de tomar como señal de falta de control la aparición de un punto fuera de los límites, considera además las tres señales adicionales mencionadas antes para el gráfico de medias.

### **2.6.4- Simulación del gráfico CUSUM**

Con el gráfico CUSUM, trabajar con observaciones estandarizadas implica llevar dos sumas acumuladas de la forma:

FIGURA 2.3- Diagrama de flujo para la simulación del gráfico de medias con señales adicionales. Tamaño de muestra n.

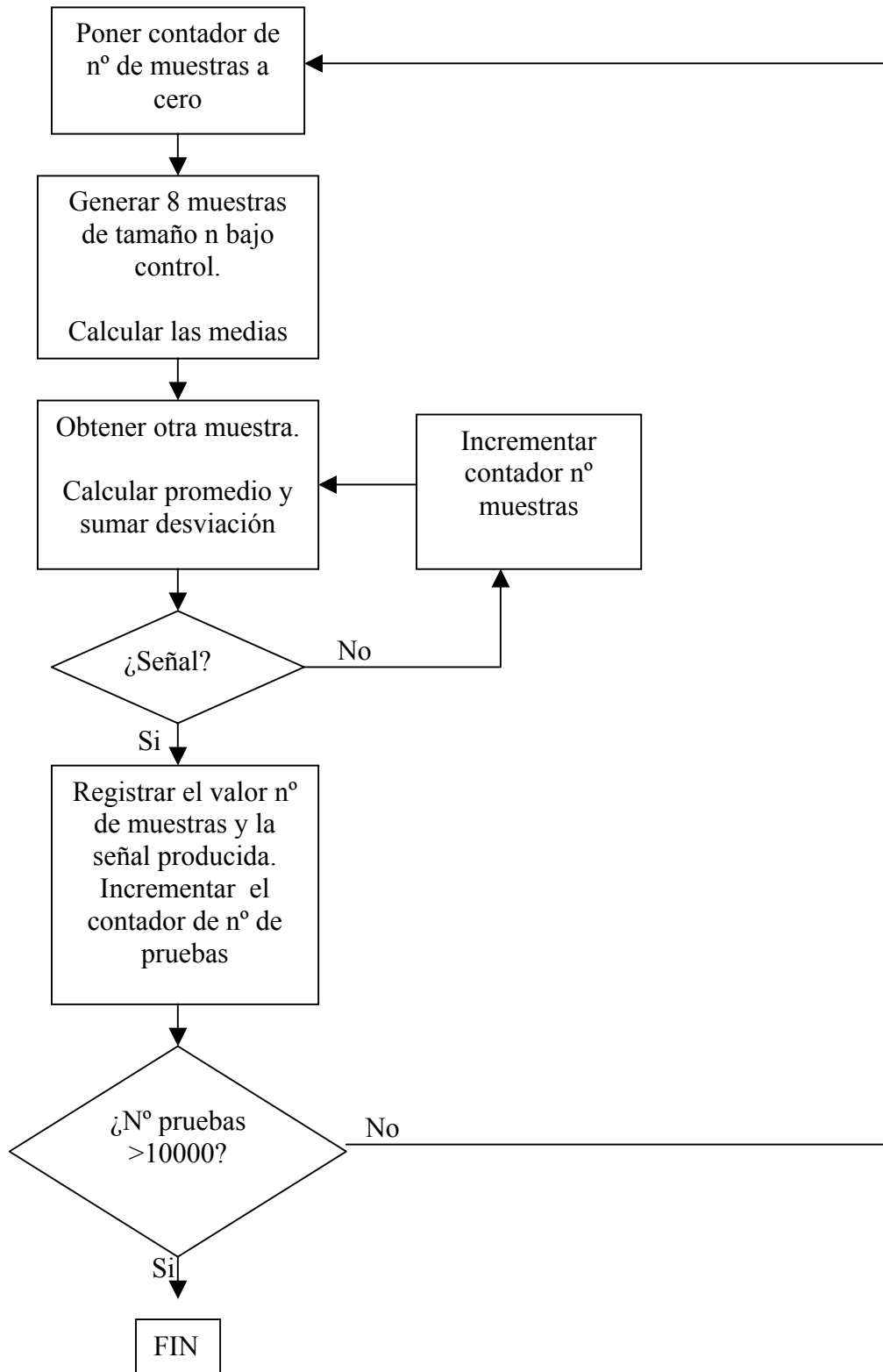
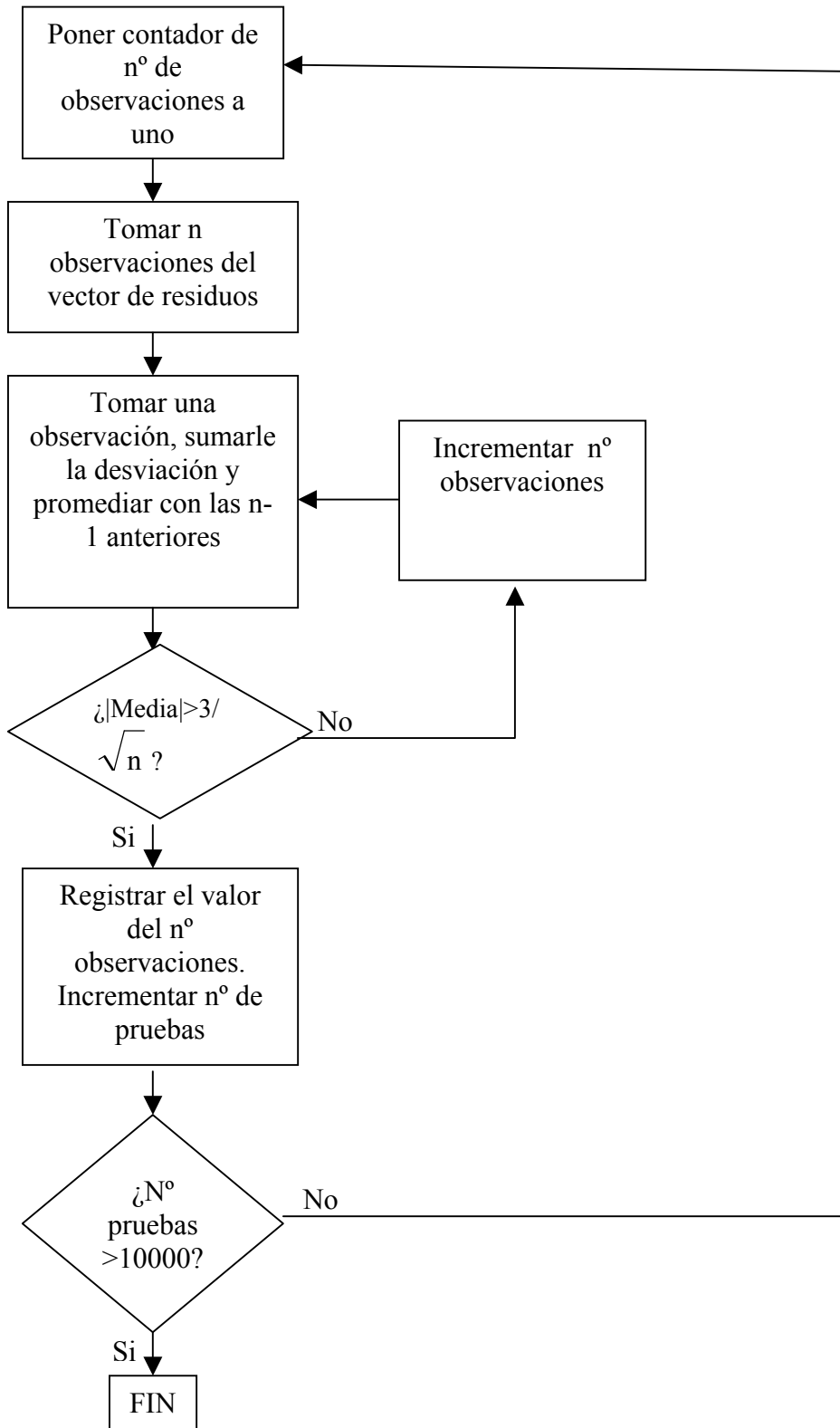


FIGURA 2.4- Diagrama de flujo para el programa de simulación del gráfico de medias móviles.





$$S_i = \max\left(0, S_{i-1} + \left(\frac{\sum_{j=1}^n \varepsilon_{ij}}{n} + \delta - \frac{k}{\sqrt{n}}\right)\right)$$

$$T_i = \max\left(0, T_{i-1} + \left(\frac{\sum_{j=1}^n \varepsilon_{ij}}{n} + \delta + \frac{k}{\sqrt{n}}\right)\right)$$

En las expresiones anteriores se toma normalmente  $S_0 = T_0 = 0$ . Tendrá lugar una señal de falta de control cuando  $T_i < -h/\sqrt{n}$  o  $S_i > h/\sqrt{n}$ , siendo  $k$  y  $h$  el valor de referencia y el intervalo de decisión respectivamente. Cuando la deriva es progresiva, las expresiones de las sumas acumuladas se modifican multiplicando la desviación por el número de muestra  $i$ . En el Anejo 1 se incluye la descripción detallada y un ejemplo de este tipo de gráfico.

En la FIGURA 2.5 se recoge el diagrama de flujo correspondiente al programa de simulación del procedimiento anterior.

En este caso, también se obtiene como resultado un vector de 10000 valores de la longitud de recorrido, que puede ser procesado para calcular el ARL, su desviación típica, etc... Como valores de entrada se requieren  $k$ ,  $h$ , tamaño de muestra y desviación en la media del proceso.

### 2.6.5- Simulación del Gráfico EWMA

Con respecto al gráfico EWMA, su diagrama de flujo se recoge en la FIGURA 2.6. En este caso el valor  $Z_t$  a comparar con los límites de control en un instante  $t$  se obtiene como una media ponderada según el parámetro  $\lambda$  entre el valor observado  $Y_{t-1}$  y el predicho  $Z_{t-1}$  para el instante  $t-1$ , según la expresión:

$$Z_t = \lambda Y_{t-1} + (1-\lambda) Z_{t-1}$$

FIGURA 2.5- Diagrama de flujo del programa de simulación para el gráfico CUSUM.

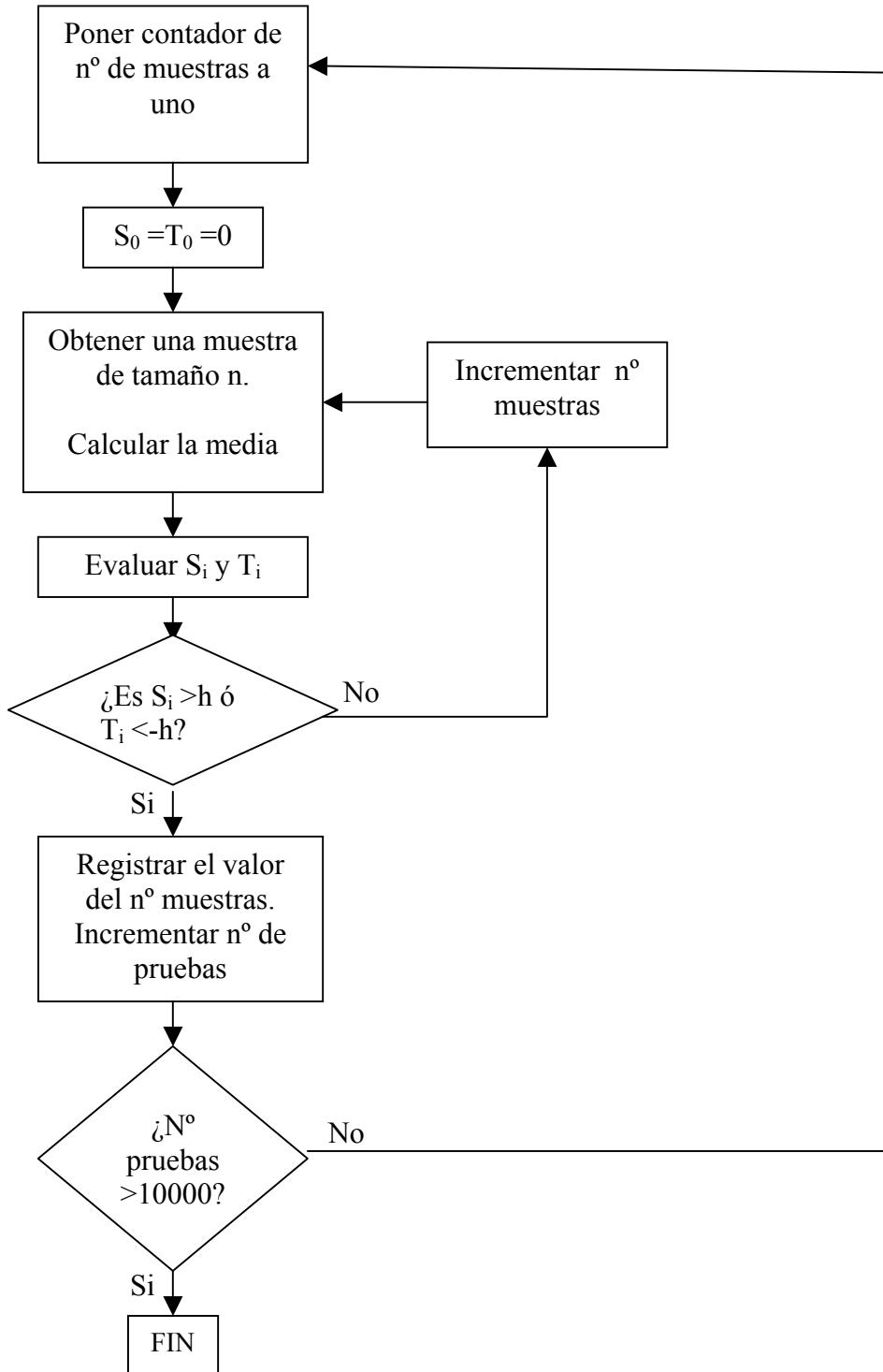
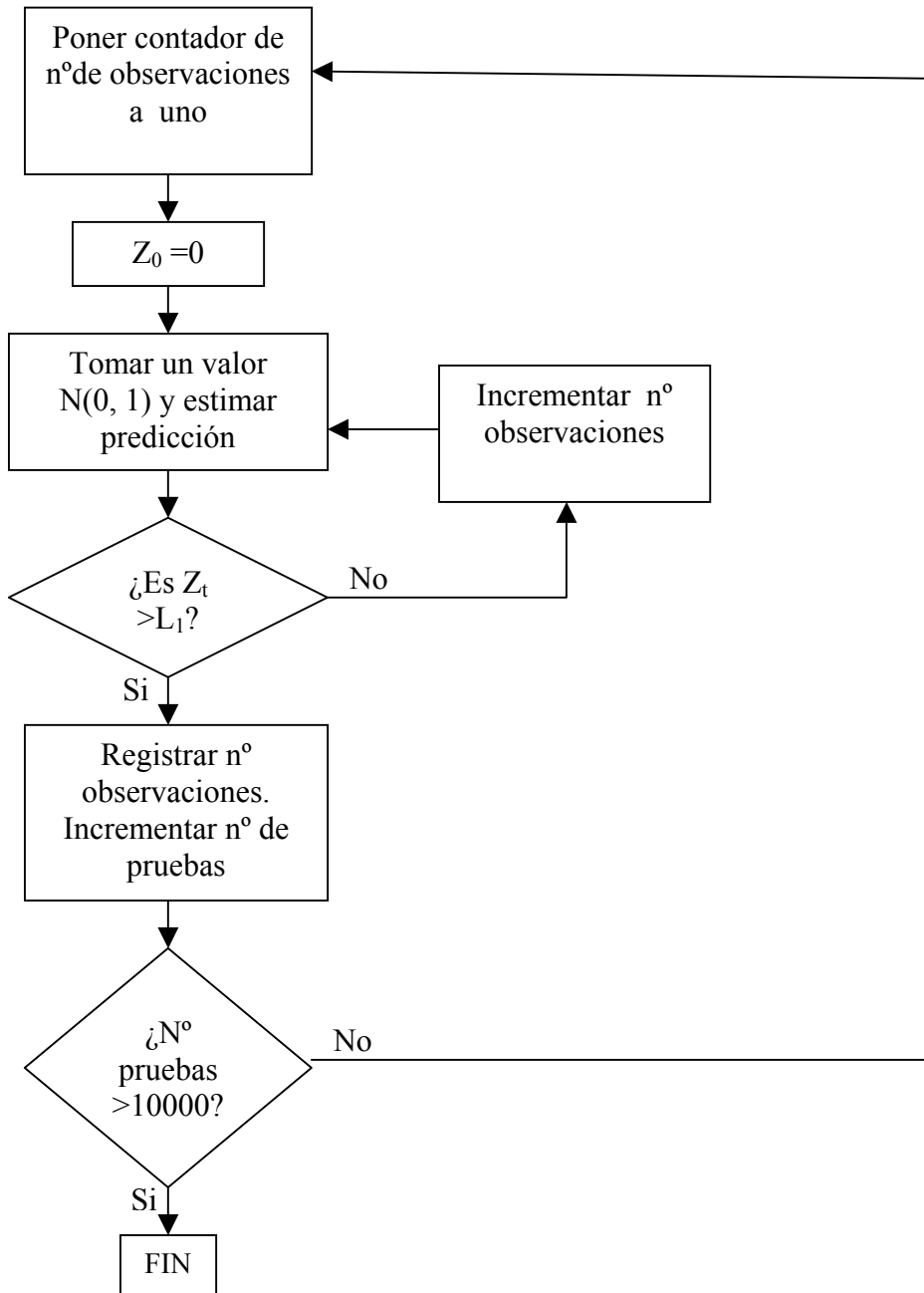


FIGURA 2.6- Diagrama de flujo para la simulación del gráfico EWMA aplicado a medidas individuales.



En este caso las observaciones individuales  $Y_i$  son valores de una  $N(0,1)$  obtenidos al azar. El estadístico  $Z_0$  para el instante inicial se suele tomar igual al nominal, que en nuestro caso es igual a 0.

La señal de falta de control ocurre cuando un valor  $Z_t$  supera los límites de control, situados en la posición:

$$0 \pm L \sqrt{\frac{\lambda}{(2-\lambda)}}$$

Cuando en vez de tratarse de valores individuales, el EWMA se aplica a medias muestrales, en la expresión anterior se considera también el tamaño de muestra.

#### **2.6.6- Simulación de observaciones correlacionadas**

Finalmente, en la última parte de la presente tesis se ha abordado el análisis de las propiedades de los gráficos CUSUM, medias móviles, EWMA y medidas individuales, cuando las observaciones están correlacionadas entre sí.

Los programas para la simulación son una modificación de los explicados antes, en el sentido siguiente. Las observaciones ya no son escogidas al azar del vector de perturbaciones aleatorias  $N(0,1)$ , sino que se generan secuencialmente según un determinado modelo ARIMA.

Las perturbaciones aleatorias se utilizan para generar la serie según se indica en el apartado 2.3. El resto del programa en cada caso es idéntico al correspondiente para observaciones independientes, por lo que no se incluye el diagrama de flujo.

El código APL de los programas antes descritos se puede encontrar en el Anejo 2.

## **CAPÍTULO 3**

### **FRECUENCIA Y TAMAÑO DE MUESTREO EN GRÁFICOS DE MEDIAS**

#### **3.1- INTRODUCCIÓN**

Cuando la característica de calidad a controlar es una variable medible en una escala continua y la producción está constituida por un gran número de unidades individuales o piezas, el procedimiento de control más difundido es el gráfico de medias y recorridos.

En concreto, el gráfico de recorridos se utiliza para controlar la dispersión de la variable de interés, cuestión que no vamos a abordar en la presente tesis, por lo que se supondrá de ahora en adelante que la desviación típica del proceso es constante. Centraremos el estudio en el gráfico de medias propuesto inicialmente por Shewhart (1931). En el Anejo 1 se incluye una explicación sobre la construcción del mismo.

Una de las cuestiones a considerar en la etapa previa a su implantación, es la determinación del plan de muestreo óptimo. En concreto, nos hemos centrado en aquellas situaciones en las que el coste total de muestreo es aproximadamente proporcional al número total de unidades muestreadas. En estos casos se plantea a menudo el dilema de elegir entre planes basados en muestras pequeñas tomadas a intervalos de tiempo cortos, frente a otros consistentes en muestras de mayor tamaño pero más espaciadas en el tiempo.

Se trata, por tanto, de analizar cómo dicha elección afecta a la potencia del gráfico, definida ésta como su capacidad para detectar rápidamente las salidas de control del proceso. Adicionalmente, este análisis permite, bajo ciertas hipótesis razonables, establecer el tamaño óptimo de muestra como función de la capacidad del proceso que se pretende controlar.

### 3.2- 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 de medias se utiliza para realizar inferencias respecto de la media de dicha distribución o posición del proceso. La FIGURA 3.1 recoge un diagrama de flujo que refleja la relación existente entre un gráfico de este tipo y un contraste de hipótesis sobre la media.

Las propiedades estadísticas de un contraste de este tipo vienen caracterizadas por sus riesgos de primera y segunda especie. En el caso concreto del gráfico de medias, el riesgo de primera especie  $\alpha$  es la probabilidad de obtener una falsa señal de falta de control, es decir un punto fuera de los límites de control pese a que la media del proceso no se ha modificado.

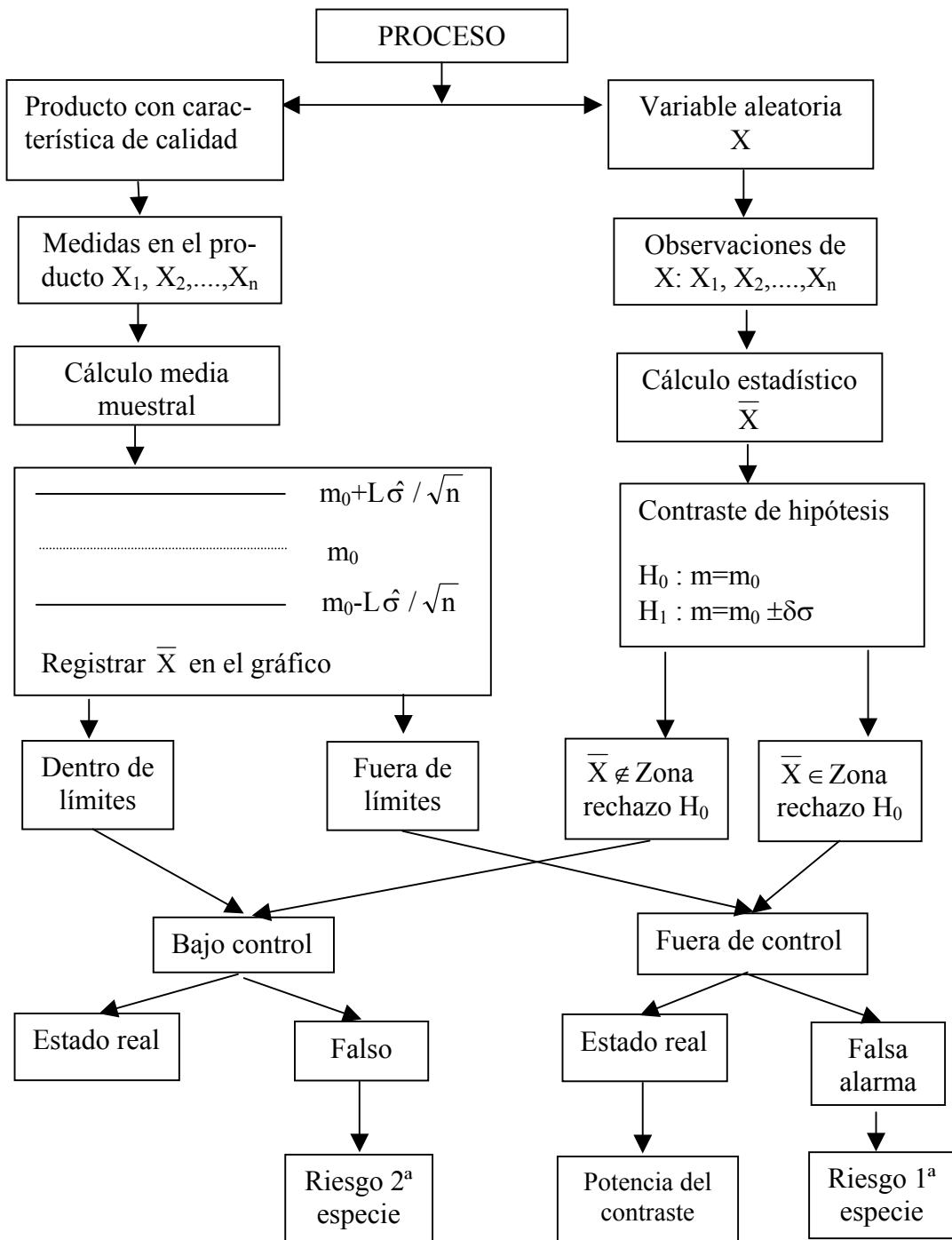
Supongamos que la característica de calidad  $X_i$  sigue una distribución aproximadamente normal, de media  $m_0$  y desviación típica  $\sigma$ . Asumiendo para el gráfico de medias una posición de los límites de control a  $\pm 3$  veces la desviación típica del parámetro a controlar respecto de  $m_0$ , y que el proceso está bajo control, el riesgo de primera especie resulta:

$$\alpha = P(\bar{X} \geq m_0 + 3\sigma/\sqrt{n} / m=m_0) + P(\bar{X} \leq m_0 - 3\sigma/\sqrt{n} / m=m_0)$$

siendo por tanto igual a 0.0027 ( 2.7% ).

El riesgo de segunda especie  $\beta$  está a su vez asociado a la probabilidad de no detectar las salidas de control reales del proceso. Teniendo en cuenta el procedimiento considerado, este riesgo es la probabilidad de que un punto no caiga fuera de los límites de control pese a que la media del proceso se ha modificado, siendo por tanto función de la magnitud de desviación producida y del tamaño de muestra  $n$ .

FIGURA 3.1- Diagrama de flujo mostrando la relación entre el gráfico de medias y un contraste de hipótesis sobre la media.



Asumiendo una desviación constante de valor  $\delta\sigma$ , el riesgo de segunda especie se puede calcular como:

$$\beta = P(\bar{X} \leq m_0 + 3\sigma/\sqrt{n} / m=m_0+\delta\sigma) + P(\bar{X} \geq m_0 - 3\sigma/\sqrt{n} / m=m_0+\delta\sigma)$$

si los límites de control se posicionan según el procedimiento estándar.

En general, para un riesgo de primera especie dado, la probabilidad de detectar una salida de control aumenta, y disminuye por tanto el riesgo de segunda especie, a medida que aumenta el tamaño  $n$  de muestra. La TABLA 3.1 refleja el valor de esta probabilidad para tamaños  $n=1, 3, 5$  y  $7$ .

Por lo tanto, si no se consideran otros parámetros de diseño del gráfico, como la frecuencia de muestreo, a mayor tamaño de muestra mayor potencia del gráfico de medias, expresada esta potencia como la probabilidad de aparición de un punto fuera de los límites de control cuando se modifica la media del proceso.

Íntimamente relacionado con los riesgos  $\alpha$  y  $\beta$  se maneja en teoría de control de procesos, como medida de la efectividad de un cierto procedimiento, el ARL o Longitud de Recorrido Medio. Tal y como se indicó en el capítulo de Metodología, este parámetro, en el caso del gráfico de medias estándar, no es más que el inverso de la probabilidad de señal de falta de control o, lo que es lo mismo, la probabilidad de aparición de un punto fuera de los límites de control. Resulta, por tanto, igual a  $1/\alpha$  cuando el proceso está bajo control, e igual a  $1/1-\beta$  (ó  $1/P(\text{señal})$ ) cuando se produce una desviación en la media del proceso.

En la TABLA 3.2 se recogen los valores del ARL para los casos considerados en la TABLA 3.1. Según se observa, para un mismo número promedio de falsas alarmas, a mayor tamaño de muestra menor valor del ARL, o lo que es lo mismo, mayor rapidez en la detección de la falta de control si en todos los procedimientos se ha empleado la misma frecuencia de muestreo.

Ahora bien, cuando se comparan planes de control con diferentes frecuencias de muestreo resulta relevante el concepto de AvD (Average Duration), 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 :

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

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

TABLA 3.1- Tablas de potencia del gráfico de medias estándar en función del tamaño de muestra  $n$  y la desviación  $\delta$ .

$\delta$ (Ud $\sigma$ )	$n=1$	$n=3$	$n=5$	$n=7$
0.00	0.002699	0.002699	0.002699	0.002699
0.25	0.003557	0.005428	0.007509	0.009704
0.50	0.006442	0.016777	0.029939	0.046766
0.75	0.012313	0.044483	0.092927	0.154889
1.00	0.022782	0.102409	0.222453	0.361574
1.25	0.040069	0.201876	0.418817	0.620652
1.50	0.066810	0.343868	0.638371	0.833635
1.75	0.105650	0.512404	0.819411	0.948456
2.00	0.158655	0.678714	0.929508	0.989033
2.25	0.226626	0.815172	0.978880	0.998426
2.50	0.308536	0.908262	0.995204	0.999849
2.75	0.401292	0.961062	0.999181	0.999999
3.00	0.500000	0.985960	0.999896	1
3.50	0.691464	0.998900	0.999999	1
4.00	0.841345	0.999999	1	1
4.50	0.933195	1	1	1
5.00	0.977250	1	1	1

Se ha visto que, con carácter general y por su relación con los riesgos de primera y segunda especie, el ARL disminuye 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  $F$  del proceso, la diferencia entre la nueva media y la existente bajo control.

TABLA 3.2- ARL del gráfico de medias estándar en función del tamaño de muestra y de la desviación.

$\delta$ (Ud $\sigma$ )	n=1	n=3	n=5	n=7
0	370.3978	370.3978	370.3978	370.3978
0.25	281.1437	184.237	133.159	103.046
0.5	155.2241	60.6883	33.4009	21.3828
0.75	81.2163	22.4803	10.7611	6.4562
1	43.8949	9.7647	4.4953	2.7657
1.25	24.9565	4.9535	2.3877	1.6112
1.5	14.9677	2.9081	1.5665	1.1996
1.75	9.4652	1.9516	1.2204	1.0543
2	6.3029	1.4734	1.0758	1.0111
2.25	4.4125	1.2267	1.0216	1.0016
2.5	3.2411	1.1010	1.0048	1.0001
2.75	2.4919	1.0405	1.0008	1
3	2.0000	1.0142	1	1
3.5	1.4462	1.0011	1	1
4	1.1886	1	1	1
4.5	1.0716	1	1	1
5	1.0233	1	1	1

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 se detecten. Las segundas ayudan a distribuir dicho esfuerzo de forma que se optimicen las propiedades estadísticas del gráfico.

El coste de las unidades defectuosas producidas dependerá del tipo de producto y proceso que se esté considerando en cada caso. Por lo tanto, centraremos nuestro estudio en el caso particular de coste de muestreo aproximadamente proporcional al número de unidades muestreadas. Bajo esta hipótesis 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 el siguiente apartado.

### **3.3- RELACIÓN DEL AvD CON EL TAMAÑO Y FRECUENCIA DE MUESTREO**

Se plantea la cuestión de elegir entre planes basados en muestras pequeñas tomadas a intervalos cortos, frente a otros consistentes en muestras de mayor tamaño más espaciadas en el tiempo. Sin pérdida de generalidad, 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. Teniendo en cuenta la simplificación asumida para el coste de muestreo, todos los procedimientos resultan en consecuencia equivalentes desde el punto de vista de coste.

El intervalo de muestreo  $f$  viene expresado en todos los casos como múltiplo de una unidad de tiempo arbitraria. Los planes a comparar son los abajo indicados. Se ha cubierto la gama de situaciones más frecuentes en la práctica industrial, en la que los tamaños de muestra no suelen superar el mayor considerado en este caso, sobre todo si la desviación típica del proceso se va a estimar a partir de los recorridos muestrales.

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:  $AvD=(ARL-0.5)xf$ , donde ARL es igual a  $1/p$ , siendo  $p$  la probabilidad de que se produzca una señal de falta de control. Como única señal de este tipo se ha considerado la aparición de un punto fuera de los límites del gráfico de medias. También se asume, de la forma habitual, que la media muestral sigue aproximadamente una distribución normal.

Con el fin de obtener procedimientos comparables en función del parámetro AvD, se ha ajustado la posición de los límites de control de forma que el tiempo promedio 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 de muestra  $n=5$ , procedimiento habitualmente recomendado, tal y como se indica en el apartado de revisión bibliográfica, con la frecuencia de muestreo indicada para el procedimiento 3. Los límites de control se han situado en la posición estándar  $m \pm 3\sigma/\sqrt{5}$ . Para este procedimiento, cuando el proceso está bajo control se tiene que  $ARL = 1/\alpha = 370.37$ , siendo  $\alpha = 2.7\%$ . 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:

$$AvD = 1850 = (ARL - 0.5)xf$$

de donde puede despejarse en cada caso ARL:

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

siendo  $p$  la probabilidad de obtener un punto fuera de los límites de control, estando el proceso bajo control. A partir de este resultado se puede calcular la posición  $m \pm L\sigma/\sqrt{n}$  de los restantes procedimientos, obteniéndose los resultados que se reflejan en la TABLA 3.3.

TABLA 3.3- Posición de los límites de control y riesgo de primera especie  $\alpha$  de los procedimientos a comparar.

PLAN	L	$\alpha$ (‰)
n=1 f=1	3.46	0.54
n=3 f=3	3.15	1.62
n=5 f=5	3.00	2.70
n=7 f=7	2.9	2.77

Si la media del proceso sufre una desviación  $\delta\sigma$ , el nuevo nivel es igual a  $m_0 + \delta\sigma$ . El ARL correspondiente a esta situación, para un procedimiento con un tamaño  $n$  de muestra y una frecuencia  $f$  de muestreo, se puede calcular tal y como se indicó en capítulo 2 como  $1/1-\beta$ . Sin embargo en este caso la posición de los límites de control varía de un procedimiento a otro.

Los resultados correspondientes para el valor del ARL en función del procedimiento y la desviación  $\delta$ , se recogen en la TABLA 3.4.

TABLA 3.4 - ARL de los procedimientos elegidos, en función de la desviación.

$\delta$ (Ud $\sigma$ )	n=1 f=1 L=3.46	n=3 f=3 L=3.15	n=5 f=5 L=3.0	n=7 f=7 L=2.9
0	1851.3	616.9	370.4	264.7
0.5	649.3	88.3	33.2	17.2
1	144.9	12.8	4.5	2.5
1.5	42.4	3.4	1.8	1.2
2	13.7	1.6	1.1	1
2.5	5.4	1.1	1	1
3	3.1	1	1	1
3.5	1.9	1	1	1
4	1.4	1	1	1

A partir de los valores del ARL recogidos en esta tabla, se puede calcular el AvD utilizando la fórmula dada anteriormente. Los resultados obtenidos para este parámetro se reflejan en la TABLA 3.5 y en la correspondiente FIGURA 3.2, en la que se utiliza escala logarítmica para los valores del AvD.

Como puede apreciarse el plan de muestreo óptimo, entendiendo como tal el que minimiza el AvD, depende de la magnitud de modificación de la media del proceso. Se observa 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 Uds. de tiempo. Por otro lado, para detectar precozmente desviaciones de magnitud grande (v.g.  $3.5\sigma$ ) la mejor solución consiste en tomar muestras de tamaño menor más frecuentemente.

TABLA 3.5 - AvD del gráfico de medias en función del procedimiento de muestreo y desviación de la media del proceso (medida en Ud. de desviación típica del proceso). Límites de control ajustados de forma que el intervalo entre falsas alarmas es el mismo en todos los casos.

$\delta$ (Ud $\sigma$ )	n=1 f=1	n=3 f=3	n=5 f=5	n=7 f=7
0	1850.0	1850.0	1850.0	1850.0
0.5	649.0	263.0	164.0	117.0
1	144.5	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

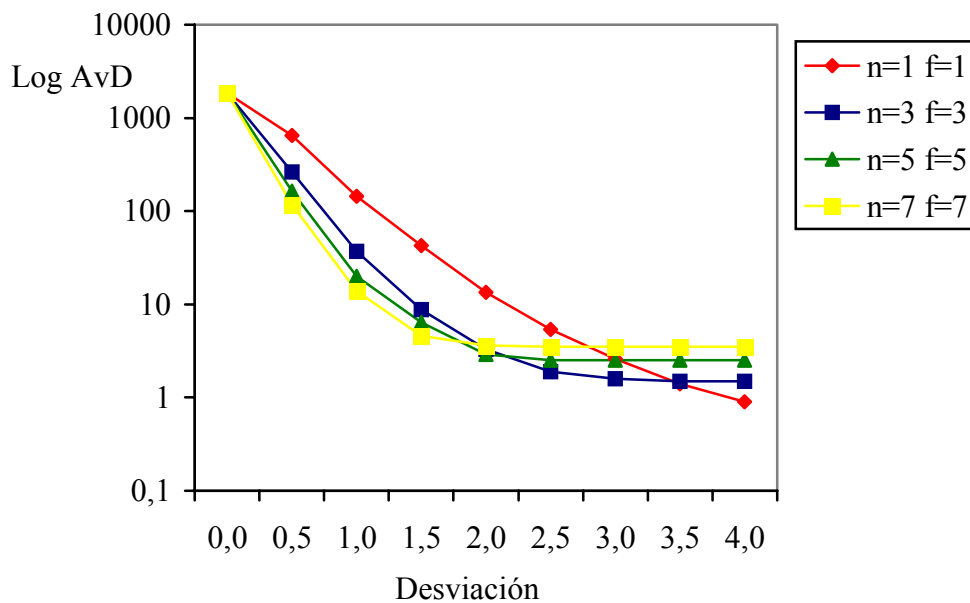
La forma de utilizar la información contenida en la tabla y el gráfico mencionados es la siguiente: si se conoce la desviación típica del proceso  $\sigma$ , una vez determinada la magnitud de desviación en la media que se considera importante, se

calcula el valor de  $\delta$  y a continuación se lee a partir de la tabla o figura, qué procedimiento resulta el óptimo.

En la práctica habitual, se desconoce normalmente el valor de  $\sigma$ , teniendo que ser estimado inicialmente a partir de los recorridos muestrales correspondientes a un cierto número de subgrupos. Cuando  $n=1$ ,  $\sigma$  se estima a partir de recorridos móviles asumiendo un cierto tamaño de agrupamiento.

En cualquier caso, se calcula el valor del recorrido medio  $\bar{R}$ , adoptando de partida un tamaño y frecuencia de muestreo según la magnitud de desviación que se considere importante. Si esta desviación es grande, el plan óptimo consiste en muestras pequeñas tomadas a intervalos relativamente cortos de tiempo. La decisión a este respecto dependerá del conocimiento previo que se tenga del proceso y de la característica a controlar. A continuación, la desviación típica  $\sigma$  se estima como  $\bar{R}/d_2$ , y a partir de ésta el valor de  $\delta$ . Acudiendo a la tabla o figura anteriores, se puede conocer el valor del AvD correspondiente al procedimiento elegido y en función del resultado, adoptar éste o probar otro distinto.

FIGURA 3.2- AvD en función de la desviación en la media y el plan de muestreo.



El plan de muestreo a adoptar depende, por tanto, del orden de magnitud de las desviaciones respecto del nominal cuya detección precoz resulte relevante. La desviación vendrá expresada en unidades de desviación típica del proceso, pudiendo ser esta última conocida o bien estimada a partir de un cierto número de subgrupos, en los que se adoptará de forma tentativa un tamaño y frecuencia de muestreo determinados.

Adicionalmente, se desprende de lo expuesto anteriormente, que el plan de muestreo está a su vez relacionado con la capacidad del proceso a controlar, ya que ésta determina la magnitud de desviación crítica, según se analiza en el siguiente apartado.

### **3.4- RELACION ENTRE EL PLAN DE MUESTREO Y LA CAPACIDAD DEL PROCESO**

El índice de capacidad potencial de un proceso  $C_p$  mide la relación existente entre la anchura de especificación fijada para la característica de calidad y la variabilidad natural del proceso, definida esta última como seis veces la desviación típica del mismo,

$$C_p = (LES - LEI) / 6F$$

En la expresión anterior LES y LEI son los límites de especificación superior e inferior para la característica en cuestión.

En un proceso centrado el valor del  $C_p$  coincide con el de otro índice, el  $C_{pk}$  o índice de capacidad real del proceso, permitiendo obtener el porcentaje de la fabricación total que se encuentra fuera de especificación, cuando el proceso está 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  $C_p$ .

Teniendo en cuenta lo anterior, se ha elaborado la TABLA 3.6, en la que se recoge el porcentaje de producto fuera de especificación, en función del  $C_p$  del proceso y de la magnitud de desviación producida en la media, expresada en unidades de  $\sigma$  del



proceso. Así, un proceso centrado con una anchura de especificaciones igual a  $4\sigma_F$ , estando bajo control presentará fuera de especificación un 4.56% del total de lo producido, asumiendo una distribución aproximadamente normal para la característica de calidad.

TABLA 3.6- Proporción de producto fuera de especificación en función de la capacidad y la desviación de la media.

$\delta$ (Ud $\sigma$ )	LES-LEI (Ud $\sigma$ )				
	3 Cp=0.50	4 0.67	6 1.00	8 1.3	9 1.5
0.00	13.36	4.56	0.27	0.001	0.0003
0.25	14.56	5.22	0.35	0.003	0.001
0.50	18.13	7.30	0.64	0.009	0.003
0.75	23.88	10.86	1.23	0.023	0.009
1.00	32.07	15.99	2.28	0.058	0.023
1.25	40.41	22.73	4.01	0.135	0.058
1.50	50.13	30.87	6.69	0.298	0.135
1.75	59.93	40.13	10.56	0.621	0.298
2.00	69.16	50.00	15.86	1.22	0.621
2.25	77.62	59.87	22.66	2.27	1.22
2.50	84.14	69.15	30.85	4.00	2.27
2.75	89.44	77.34	40.13	6.68	4.00
3.00	93.32	84.13	50.00	10.56	6.68
3.25	95.99	89.44	59.87	15.86	10.56
3.50	97.72	93.32	69.15	22.66	15.86
4.00	99.38	97.72	84.13	40.13	30.85
4.50	99.86	99.38	93.32	59.87	50.00
5.00	99.97	99.86	97.72	77.34	69.15

A partir de esta tabla y de los resultados obtenidos para el AvD, es posible establecer el plan de muestreo óptimo según la capacidad del proceso.

Por ejemplo, supongamos un proceso con anchura de especificaciones igual a  $6\sigma$  ( $C_p=1$ ) en el que además se desea detectar lo más rápidamente posible una situación en la que aproximadamente el 10% de lo producido esté fuera de especificaciones. Según se observa en la TABLA 3.6, esta situación se da cuando la media del proceso sufre una desviación igual a  $1.75F$ . De la FIGURA 3.2 o la TABLA 3.5, se concluye que de los procedimientos estudiados el que detecta esta situación más rápidamente consiste en tomar muestras de tamaño 7 cada 7 unidades de tiempo.

En general, 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 se va a asumir que se trata de minimizar el tiempo promedio, o AvD, que se tarda en detectar una desviación respecto del valor nominal que produce un porcentaje de defectuosos igual o superior al 50%.

En este caso concreto, la desviación  $\delta$  que produce este porcentaje será igual a  $3xC_p$ , tal y como se observa también a partir de la TABLA 3.6. Teniendo esto en cuenta, se ha elaborado la TABLA 3.7, y a partir de ella la FIGURA 3.3, que representa el tamaño de muestra óptimo para detectar esta situación, asumiendo constancia en el número total de unidades muestreadas y en función de la capacidad  $C_p$  del proceso.

Para los planes de muestreo no considerados anteriormente, los límites de control han sido ajustados en las posiciones siguientes:

$$\begin{aligned} n=2 \quad f=2 & \quad m \pm 3.27\sigma/\sqrt{2} \\ n=4 \quad f=4 & \quad m \pm 3.07\sigma/\sqrt{4} \\ n=6 \quad f=6 & \quad m \pm 2.94\sigma/\sqrt{6} \\ n=8 \quad f=8 & \quad m \pm 2.85\sigma/\sqrt{8} \end{aligned}$$

En la tabla y figura mencionadas se aprecia que cuando el proceso opera con un índice de capacidad  $C_p$  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  $C_p$  bajas el procedimiento óptimo consiste en tomar muestras de tamaño grande disminuyendo la frecuencia de muestreo.

TABLA 3.7- Plan de muestreo óptimo para detectar 50% defectuosos, en función del  $C_p$  del proceso.

$C_p$	Desviación (Ud $\sigma$ )	% Antes desviación	% Después desviación	n y f óptimos
0.33	1	31.74	52.27	n=8 f=8
0.42	1.25	21.12	50.62	n=8 f=8
0.50	1.5	13.36	50.13	n=7 f=7
0.58	1.75	8.02	50.02	n=6 f=6
0.67	2	4.56	50.00	n=5 f=5
0.75	2.25	2.44	50.00	n=4 f=4
0.83	2.5	1.24	50.00	n=3 f=3
1.00	3	0.27	50.00	n=2 f=2
1.17	3.5	0.046	50.00	n=2 f=2
1.33	4	0.003	50.00	n=1 f=1
1.42	4.25	0.001	50.00	n=1 f=1
1.50	4.5	0.0003	50.00	n=1 f=1

El valor  $n=5$ , recomendado frecuentemente en la práctica, resulta adecuado para la rápida detección de desviaciones que conducen a un 50% de defectuosos, cuando la capacidad del proceso está comprendida entre 0.6 y 0.8. Sin embargo, si lo que se desea detectar rápidamente es una situación en la que el 10% esté fuera de especificaciones, el tamaño de muestra  $n=5$  resulta óptimo si el  $C_p$  está comprendido entre 1 y 1.2.

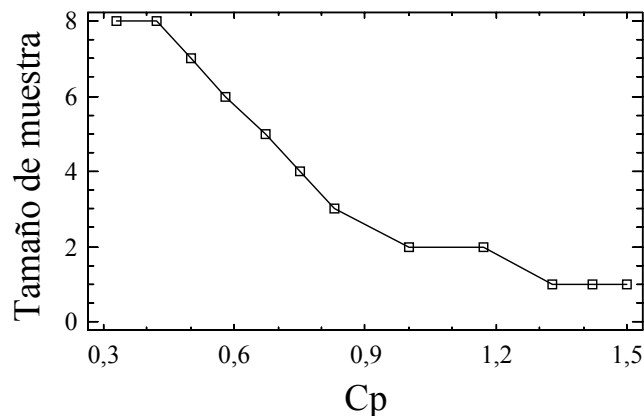
En algunos casos existe para la característica de calidad una sola especificación, superior o inferior y el índice de capacidad del proceso,  $C_{pi}$  o  $C_{ps}$ , depende de la relación entre la distancia del proceso a la especificación y la mitad de la variabilidad natural del proceso,  $3\sigma$ .

Sin pérdida de generalidad, se va a suponer que existe un límite de especificación inferior. El Cpi se calculará como:

$$C_{pi} = (m - LEI)/3\sigma$$

Asumiendo que se desea detectar un desplazamiento de la media del proceso que lleve a un 20% o más de artículos defectuosos, se ha elaborado la TABLA 3.8, en función del Cpi antes de producirse la desviación.

FIGURA 3.3- Plan de muestreo óptimo en función de Cp



Según se observa, las conclusiones son las mismas que en el caso de especificación bilateral. Cuando el proceso opera con un Cpi alto (>1.25) el plan de muestreo óptimo consiste en tomar muestras de tamaño n=1 frecuentemente, sin embargo para procesos con capacidad baja, el tamaño de muestra óptimo es mayor, con muestras más espaciadas en el tiempo.

Por último, cabe señalar que el resultado obtenido coincide en cierto sentido con la propuesta de control óptimo de Taguchi (1981), consistente en tomar muestras de tamaño pequeño a intervalos cortos de tiempo. Esto no es sorprendente dada la en general elevada capacidad de los procesos en la industria japonesa. Ahora bien, tal y como se ha mostrado en este capítulo, cuando la capacidad del proceso es baja el procedimiento de muestreo óptimo consiste en tomar muestras de mayor tamaño, más espaciadas en el tiempo.

TABLA 3.8- Relación entre capacidad y plan de muestreo con especificación unilateral.

C <sub>pi</sub>	Desviación (Ud $\sigma$ )	Procedimiento óptimo
0.33	0.25	n=8 f=8
0.42	0.50	n=8 f=8
0.50	0.75	n=7 f=7
0.58	1.00	n=6 f=6
0.67	1.25	n=5 f=5
0.75	1.50	n=4 f=4
0.83	1.75	n=3 f=3
0.92	2.00	n=3 f=3
1.00	2.25	n=2 f=2
1.08	2.50	n=2 f=2
1.17	2.75	n=2 f=2
1.25	3.00	n=1 f=1

## CAPÍTULO 4

### SEÑALES DE FALTA DE CONTROL EN GRÁFICOS DE MEDIAS

#### 4.1- INTRODUCCIÓN

Los gráficos de control de Shewhart constituyen hoy en día una de las herramientas más utilizadas para control on-line de procesos industriales. El objetivo de los mismos es detectar la existencia de causas especiales de variabilidad y ayudar a su identificación y posterior eliminación, manteniendo así el proceso bajo control.

El tipo de gráfico más ampliamente utilizado, cuando la característica de calidad toma valores en una escala continua, es el gráfico de Shewhart para la media o gráfico de medias muestrales. Básicamente, consiste en ir registrando el valor medio de cada muestra y considerar señal de falta de control un punto fuera de los límites de control. Estos límites, en el caso estándar, se sitúan a una distancia  $3\sigma/\sqrt{n}$  por encima y por debajo de la media del proceso, siendo  $n$  el tamaño de muestra y  $\sigma$  la desviación típica del proceso.

Este procedimiento presenta el inconveniente de ser poco efectivo en la detección de desviaciones de baja magnitud en la media del proceso. Con el fin de mejorarlo en este sentido, se introducen en la práctica habitual señales de falta de control adicionales, basadas en tendencias o rachas que pueden aparecer en el gráfico. De esta forma, no se considera sólo la posición del último punto graficado sino también la información contenida en un determinado número de muestras anteriores.

El introducir nuevas señales de falta de control en el gráfico de medias supone, en primer lugar, una complejidad de manejo mucho mayor. A nivel práctico se debe estudiar si la mejora que se produce en sensibilidad compensa la dificultad que podría suponer su utilización en planta. Por otro lado, las propiedades estadísticas del gráfico

cambian, aumentando al mismo tiempo que la sensibilidad a bajas desviaciones, la frecuencia relativa de falsas alarmas.

Tal y como se señala en la revisión bibliográfica, los gráficos de control de sumas acumuladas, gráficos CUSUM, se caracterizan, en general, por una mayor efectividad para la detección de desviaciones de poca magnitud, que los gráficos de medias estándar. Si bien esto último depende de los parámetros adoptados en su diseño, la mayoría de las referencias mencionadas coinciden en señalar que el gráfico CUSUM presenta un ARL menor que el del gráfico de medias estándar, cuando la desviación en la media es constante y pequeña. Sin embargo, no se ha abordado el estudio comparativo de estos dos gráficos, cuando además de la señal consistente en un punto fuera de los límites, se incluyen las tres señales adicionales definidas en apartados anteriores.

Tiene interés el estudio de esta última cuestión, puesto que el gráfico CUSUM y el gráfico de medias con señales adicionales difieren considerablemente en cuanto a la complejidad de diseño e implantación. Adicionalmente, con el fin de cuantificar la variación de sus propiedades, se incluye en este estudio el gráfico de medias estándar, pero ajustando sus límites para que los ARL bajo control coincidan.

Se aborda por tanto, mediante el recurso a técnicas de simulación el estudio comparativo de los tres procedimientos antes señalados: gráfico de medias estándar, gráfico de medias con señales adicionales y gráfico CUSUM para la media. Se pretende analizar su comportamiento en los casos de desviación constante y progresiva en la media del proceso, simulando distintas situaciones que difieren en la magnitud y naturaleza de esta desviación, que en general expresaremos en unidades de desviación típica del proceso.

Por otro lado, se consideran también distintos tamaños de muestra. Con el fin de cubrir la gama de situaciones más corrientes en la práctica, se han tomado  $n=1, 3, 5$  y  $7$ . El caso  $n=1$ , corresponde con el gráfico de medidas individuales, en el que también se recomienda habitualmente la inclusión de señales adicionales de falta de control.

El parámetro utilizado como medida de la efectividad de cada gráfico, para establecer comparaciones entre los mismos, es el ARL, valor medio del número de muestras tomadas desde que el proceso se sale de control hasta que se produce señal de falta de control. En el capítulo de metodología se incluyó una discusión sobre este parámetro y procedimientos para su estimación.

Con el fin de precisar además la importancia de las señales adicionales de falta de control se obtiene también la frecuencia relativa de aparición de cada una bajo diferentes hipótesis respecto al tamaño de la muestra y a la naturaleza de la salida de control. En relación con esto último se ha considerado también para su interpretación, su frecuencia cuando la desviación se produce en la dispersión de la característica de calidad.

#### 4.2- CONSIDERACIONES PREVIAS

Diversos autores que se recogen en la revisión bibliográfica, han estudiado el comportamiento de los gráficos de medias cuando se utilizan señales de falta de control adicionales a la propuesta por Shewhart. Estas señales pueden resumirse básicamente de la siguiente forma: tiene lugar una señal de falta de control cuando  $k$  de las  $p$  últimas medias muestrales forman una racha de determinadas características.

Existen distintas propuestas de señales adicionales. Sin embargo, en el contexto de la industria española, las señales de falta de control más utilizadas son las siguientes:

**Señal 1:** Punto fuera de los límites  $m \pm 3 \sigma / \sqrt{n}$ .

**Señal 2:** Dos de tres puntos seguidos en la zona comprendida entre los llamados límites de aviso  $m \pm 2 \sigma / \sqrt{n}$  y los de control  $m \pm 3 \sigma / \sqrt{n}$ .

**Señal 3:** Siete puntos seguidos por encima o por debajo de la línea central.

**Señal 4:** Racha ascendente o descendente de siete puntos seguidos.



El estudio de las propiedades del gráfico de medias, cuando se incluye este conjunto de señales en particular, tiene interés debido a que no se encuentra resuelta esta cuestión en la bibliografía sobre el tema. Por otro lado, las señales aquí definidas difieren de otras propuestas, para las que sí se ha abordado la cuestión, por lo que en principio cabe esperar resultados distintos. Así, en ninguna de las referencias hemos encontrado una señal similar a la racha ascendente o descendente de siete puntos seguidos.

En las salidas de control simuladas, se consideran las dos situaciones de desviación constante y progresiva en la media del proceso. En el primer caso, una vez producida una desviación de magnitud  $\delta\sigma$ , la media del proceso se mantiene fija a un nuevo nivel  $m+\delta\sigma$  y en el segundo la nueva media es igual a  $m+\delta_i\sigma$ , siendo  $i=1,2,\dots$  el número de muestra desde que el proceso se salió de control.

Tal y como se indicó en el apartado de metodología, el cálculo analítico del ARL correspondiente a un gráfico de medias con las cuatro señales antes definidas resulta una tarea inabordable. En primer lugar, estas señales no son independientes entre sí. Adicionalmente, están definidas de forma que dependen a su vez de la información contenida en muestras anteriores. Esto imposibilita aplicar el sencillo enfoque analítico que en apartados anteriores se propone para el cálculo del ARL del gráfico estándar, puesto que la probabilidad de señal de falta de control depende del instante de tiempo en que nos encontremos.

Por otro lado, algunos de los métodos aproximados para otras combinaciones de señales de falta de control resultan inabordables si se incluye la señal de racha ascendente o descendente de varios puntos seguidos.

Teniendo en cuenta lo anterior, el estudio de los ARL del gráfico de medias con las cuatro señales antes definidas se ha realizado mediante el recurso a técnicas de simulación en ordenador. Básicamente, el procedimiento ha consistido en generar una serie de valores con distribución normal, asumir un determinado estado en el proceso y constatar el número de muestras a tomar hasta que se produce alguna de las cuatro

señales de falta de control. El diagrama de flujo del procedimiento se incluye en el apartado de metodología, y el programa APL correspondiente en el Anejo 2.

Con respecto a los gráficos CUSUM, el cálculo del ARL para una desviación constante o progresiva y un determinado tamaño de muestra, se ha realizado también mediante técnicas de simulación en ordenador. Con el fin de estimar con precisión suficiente los ARL de ambos procedimientos, gráfico de medias con señales adicionales y gráfico CUSUM, el número de simulaciones se ha tomado igual a 10000, para cada desviación de la media y tamaño muestral. Se discute también la elección de este número de repeticiones en el apartado de metodología.

El cálculo del ARL del gráfico de medias estándar con una señal de falta de control, un punto fuera de los límites, se ha realizado en el caso de desviación constante de forma analítica. El ARL correspondiente a esta situación para un procedimiento consistente en tomar muestras de tamaño  $n$  es igual a  $1/p$ , siendo  $p$  la probabilidad de que la media muestral caiga fuera de los límites  $m \pm h\sigma/\sqrt{n}$ , expresión que se puede evaluar fácilmente asumiendo distribución normal para la media muestral.

Adicionalmente, en el caso del gráfico de medias con las señales de falta de control antes definidas, se ha obtenido la frecuencia relativa con la que aparece cada una de ellas bajo las distintas hipótesis consideradas.

#### **4.3- DISEÑO DE LOS GRÁFICOS DE CONTROL A COMPARAR**

Para que sean comparables el gráfico de medias estándar, el gráfico de medias con las señales adicionales de falta de control y el gráfico CUSUM a distintas desviaciones y tamaños de muestra, se deben ajustar los parámetros de estos procedimientos con el fin de tener en los tres gráficos la misma frecuencia relativa de falsas alarmas. Esto es equivalente a utilizar gráficos con el mismo ARL bajo control.

La frecuencia relativa de falsas alarmas en el procedimiento de medias estándar, coincide con el riesgo de primera especie asociado a su operación.

Por otro lado, en el gráfico CUSUM y gráfico de medias con señales adicionales, esta frecuencia relativa es una medida aproximada de la cantidad de falsas alarmas a esperar en promedio cuando los parámetros de diseño toman un determinado valor. Se debe tener en cuenta que el riesgo de primera especie asociado a la operación de estos dos últimos gráficos no es constante, y depende del instante de tiempo en que nos encontremos, por lo que resulta inadecuado utilizarlo para definir las propiedades de estos dos gráficos. En su lugar, resulta más conveniente estudiar sus características operativas en función del ARL.

No es posible ajustar directamente los parámetros de estos tres gráficos fijando de antemano una determinada frecuencia de falsas alarmas  $\alpha$ , o ARL bajo control. Para el gráfico de medias estándar, dado un ARL bajo control, el riesgo de primera especie sería  $\alpha=1/ARL$ , y a partir de este se puede calcular fácilmente la posición de los límites de control correspondientes.

Con respecto al gráfico CUSUM, si se desea un ARL bajo control determinado, se puede acudir a las tablas o ábacos disponibles en la bibliografía sobre el tema, para fijar el valor de referencia  $k$  y del intervalo de decisión  $h$ . Sin embargo, para el gráfico de medias con señales adicionales, no existe forma analítica de determinar la posición de los límites correspondientes a un ARL bajo control determinado.

En consecuencia, se ha procedido de la siguiente manera. Se han tomado los límites para el gráfico de medias con señales adicionales en la posición estándar,  $m_0 \pm 3\sigma/\sqrt{n}$ , siendo  $n$  el tamaño de muestra y se ha calculado mediante simulación el ARL bajo control correspondiente a este procedimiento. A continuación, se han fijado los límites del gráfico de medias estándar y los parámetros  $k$  y  $h$  del CUSUM para que sus ARL bajo control coincidan respectivamente con este valor estimado.

Mediante la técnica de simulación anteriormente descrita, se obtuvo que el ARL bajo control del gráfico de medias con las señales adicionales definidas es aproximadamente igual a 60. Esto supone que en promedio el intervalo entre dos falsas alarmas consecutivas es de 60 muestras. Dicho de otra forma, aproximadamente 1 de cada 60 observaciones registradas en el gráfico cuando el proceso está bajo control resultará en una falsa alarma.. Si se desea expresar la frecuencia relativa de falsas alarmas como tanto por uno, se tiene que esta frecuencia es  $\alpha \approx 0.017$ .

Este primer resultado tiene implicaciones importantes. En el gráfico de medias estándar, con una sola señal de falta de control ( punto fuera de los límites) y límites de control a  $m_0 \pm 3\sigma/\sqrt{n}$ , la frecuencia relativa de falsas alarmas o, en este caso riesgo de primera especie, resulta igual a 0.0035, lo que supone un ARL bajo control igual a 370.

Por tanto, la introducción de señales adicionales de falta de control aumenta considerablemente la frecuencia relativa de falsas alarmas y esto implica una disminución notable en el ARL bajo control. Así, con señales adicionales la frecuencia de falsas alarmas es seis veces mayor que en el procedimiento estándar.

Teniendo en cuenta lo anterior, a falta de investigar con más detalle la potencia del gráfico para detectar las salidas de control reales del proceso, la mayor efectividad frente a desviaciones de poca magnitud que pueda suponer la consideración de señales adicionales en el gráfico de medias se consigue, en parte al menos, a costa de incrementar de forma notable la cantidad de falsas alarmas.

Los parámetros del gráfico de medias estándar y el CUSUM, se han ajustado con el fin de tener un ARL bajo control aproximadamente igual a 60. Para el gráfico de medias estándar, se trata de elegir el valor L de forma que :

$$P(|\bar{X} - m_0| > L \frac{\sigma}{\sqrt{n}} / m = m_0) = 0.017$$

de donde se deduce, asumiendo distribución normal para la media muestral, que  $L=2.39$ . Por lo tanto, los límites en este caso se sitúan en  $m_0 \pm 2.39\sigma/\sqrt{n}$ .

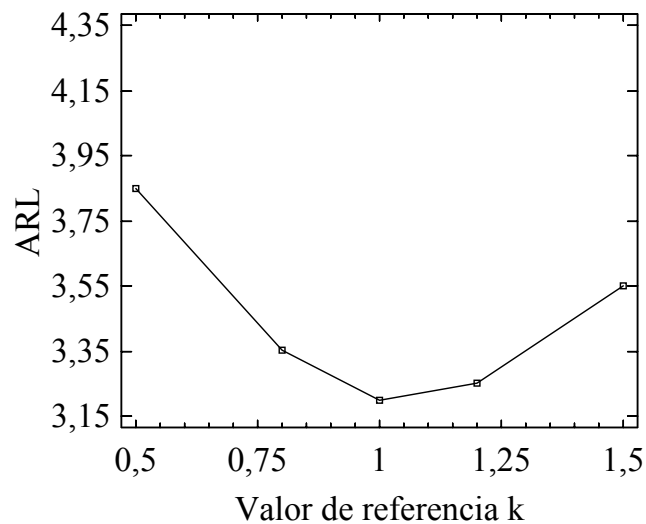
En cuanto al gráfico CUSUM, existen distintas combinaciones de  $k$  y  $h$  para un mismo ARL bajo control. Sin embargo, si se desea diseñar un gráfico especialmente sensible frente a desviaciones de poca magnitud, el valor de  $k$  debe ser pequeño, puesto que a medida que éste aumenta, las propiedades del gráfico CUSUM se aproximan más a las del gráfico de medias estándar. En general, se recomienda para la elección de los dos parámetros fijar primero el ARL bajo control, tomar  $k=\delta/2$ , siendo  $\delta$  la magnitud de desviación cuya detección precoz resulta relevante, y a partir de algún nomograma (v.g. Goel y Wu, 1972) o tabla, fijar el  $h$  que para ese  $k$  proporciona el ARL bajo control deseado.

En el caso particular que nos ocupa, no tiene demasiado sentido la consideración del valor  $\delta$  antes mencionado, ya que nuestro objetivo no se centra en un único valor de la desviación, sino en una gama más o menos amplia de valores para la misma. Sin embargo, resulta relevante para la elección de los parámetros del CUSUM con ARL bajo control aproximadamente 60, la regla práctica  $k=\delta/2$ . Este resultado, tal y como se explicó en la revisión bibliográfica, tiene importancia porque al comparar las propiedades de gráficos CUSUM con un mismo ARL bajo control, el diseñado con  $k=\delta/2$  presenta un ARL mínimo para una desviación  $\delta$  en la media.

Se ha realizado una comprobación de esto último mediante simulación, tomando varios gráficos CUSUM con un ARL bajo control aproximadamente igual a 370. En la FIGURA 4.1 se recoge el valor del ARL de los mismos cuando la desviación en la media es igual a  $2\sigma/\sqrt{n}$ , en función del valor de  $k$ , expresado también en Uds de  $\sigma/\sqrt{n}$ . Se observa que el gráfico diseñado con un valor de referencia  $k=\sigma/\sqrt{n}$  (es decir,  $k=\delta/2$ ) presenta un ARL mínimo cuando se produce una desviación de esta magnitud.

Junto a la recomendación anterior, en algunas referencias bibliográficas se considera el valor  $k=0.5\sigma/\sqrt{n}$  como estándar para el gráfico CUSUM. Teniendo en cuenta lo antes visto, si se elige  $k=0.5\sigma/\sqrt{n}$ , el gráfico CUSUM diseñado presentará, para un ARL bajo control dado, un ARL mínimo cuando la desviación en la media es igual a  $\sigma/\sqrt{n}$ .

FIGURA 4.1- ARL de gráficos CUSUM para desviación en la media igual a  $2\sigma/\sqrt{n}$  en función del valor de k. ARL bajo control aproximadamente igual a 370.



Si se desea comparar gráficos de control con el fin de determinar cuál resulta más efectivo para desviaciones bajas en la media del proceso, no interesa diseñar el CUSUM correspondiente con un valor de k alto, puesto que se pierde la ventaja de este gráfico en esas situaciones. Por otro lado, un valor de k excesivamente pequeño implica una potencia baja frente a desviaciones de magnitud media. La recomendación  $k=0.5\sigma/\sqrt{n}$  puede resultar en un gráfico CUSUM con propiedades intermedias entre estos dos casos extremos.

Por tanto, el gráfico CUSUM a considerar en nuestro análisis se va a diseñar con el valor de referencia estándar  $k = 0.5\sigma/\sqrt{n}$  y un ARL bajo control aproximadamente igual a 60. Con estos dos valores, se ha consultado el nomograma proporcionado por Goel y Wu (1972), resultando el valor de  $h = 3.04 \sigma/\sqrt{n}$ . Este nomograma se utiliza para el diseño de gráficos CUSUM para la detección de desviaciones de la media en un sólo sentido ( one-sided ).

La expresión (Page, 1954):

$$\frac{1}{\text{ARL}} = \frac{1}{\text{ARL}_1} + \frac{1}{\text{ARL}_2}$$

permite calcular el ARL de un gráfico CUSUM diseñado para detectar desviaciones en la media tanto positivas como negativas, en función de  $\text{ARL}_1$ , que es el ARL del procedimiento sólo para desviaciones positivas, y de  $\text{ARL}_2$  que es el correspondiente pero para desviaciones negativas. Si se va a diseñar el CUSUM de forma que resulte simétrico, bajo control se cumple que  $\text{ARL}_1 = \text{ARL}_2$ , con lo que  $\text{ARL} = \text{ARL}_1/2 = \text{ARL}_2/2$ . Como se desea un  $\text{ARL} = 60$ , se ha consultado el nomograma de Goel y Wu con el valor 120 y con  $k = 0.5\sigma/\sqrt{n}$ .

Teniendo en cuenta lo anterior, los gráficos a comparar son los siguientes:

**G1:** Gráfico de medias con una sola señal de falta de control, un punto fuera de los límites situados en  $m_0 \pm 2.39 \sigma/\sqrt{n}$ .

**G2:** Gráfico X con las cuatro señales antes definidas y límites en  $m_0 \pm 3 \sigma/\sqrt{n}$ .

**G3:** Gráfico CUSUM con  $k = 0.5 \sigma/\sqrt{n}$  y  $h = 3.04 \sigma/\sqrt{n}$ .

En las tablas y figuras de esta tesis nos referiremos a estos tres gráficos mediante las siglas **G1**, **G2** y **G3**.

#### **4.4- EFECTO DE LAS SEÑALES ADICIONALES SOBRE EL ARL**

Tal y como se indica en el apartado anterior, el primer efecto que tiene la inclusión de señales adicionales sobre el ARL del gráfico resultante, es una reducción considerable de su valor estando bajo control. Esto implica en la práctica, que operando en condiciones normales con el proceso bajo control, el número de falsas alarmas es excesivamente grande. Por lo tanto, a la mayor complejidad debida a la utilización de estas señales hay que sumarle el coste adicional de sobreajustar el proceso.

Tiene por tanto sentido estudiar qué ocurre con las propiedades de este procedimiento, cuando se produce alguna situación de falta de control en el proceso. Si la inclusión de señales adicionales no resulta en una mayor efectividad frente a desviaciones de poca magnitud, en comparación con el gráfico de medias estándar o el CUSUM, los inconvenientes mencionados en el párrafo anterior harán que no sea recomendable en general la utilización de las señales adicionales. Si ocurre lo contrario, se debe considerar en la fase previa a la implantación de este gráfico si el incremento de potencia compensa la mayor complejidad y mayor cantidad de falsas alarmas innatas al mismo.

La efectividad del gráfico de medias con señales adicionales se ha estudiado según el procedimiento de simulación descrito en el apartado de metodología. El parámetro utilizado en cada caso como medida de la potencia del gráfico es el ARL. Las simulaciones se han realizado para distintos tamaños de muestra, que con el fin de cubrir la gama de situaciones más corrientes en la práctica, han sido  $n=1$ , 3, 5 y 7. El valor  $n=1$  corresponde con el gráfico de medidas individuales. En cada simulación se ha asumido de partida una determinada desviación en la media, expresada en unidades de desviación típica del proceso.

#### **4.4.1- Desviación constante en la media**

En la TABLA 4.1 y FIGURA 4.2 se recogen los valores del ARL de los tres procedimientos a comparar, en función de la magnitud de desviación constante de la media del proceso. En particular, se contempla el caso, muy frecuente en la práctica, de tamaño de muestra  $n=5$ . En la figura se ha tomado escala logarítmica para los valores del ARL.

En la tabla y figura mencionadas se observa que el procedimiento de control óptimo, entendiendo como tal el que minimiza el ARL, depende de la magnitud de desviación de la media del proceso.

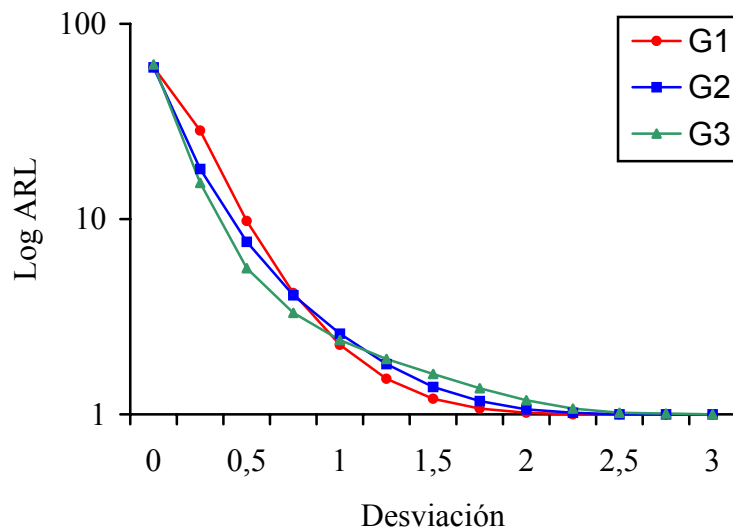


Cuando el tamaño de muestra es igual a cinco, el utilizar señales adicionales de falta de control en el gráfico de medias mejora su efectividad respecto al gráfico de medias estándar con idéntico número de falsas alarmas para desviaciones inferiores a  $0.75\sigma$ , lo que se traduce en valores del ARL menores que los correspondientes al gráfico de medias estándar. A partir de esa desviación, este último procedimiento resulta más sensible frente a cualquier salida de control en el proceso.

TABLA 4.1- ARL del gráfico de medias estándar, gráfico de medias con señales adicionales y gráfico CUSUM, en función de la magnitud de desviación constante de la media del proceso. Tamaño de muestra  $n=5$ .

$\delta$ (Ud $\sigma$ )	G1		G2		G3	
	ARL	ARL	$S_{ARL}$	ARL	$S_{ARL}$	
0.00	60.00	59.89	0.538	61.97	0.587	
0.25	28.41	18.05	0.157	15.33	0.122	
0.50	9.78	7.65	0.053	5.60	0.052	
0.75	4.18	4.07	0.025	3.31	0.031	
1.00	2.27	2.59	0.015	2.40	0.024	
1.25	1.52	1.81	0.009	1.92	0.019	
1.50	1.20	1.38	0.006	1.61	0.016	
1.75	1.07	1.17	0.004	1.36	0.005	
2.00	1.02	1.06	0.002	1.18	0.004	
2.25	1.00	1.02	0.001	1.07	0.002	
2.50	1.00	1.00	0.000	1.02	0.001	
2.75	1.00	1.00	0.000	1.01	0.000	
3.00	1.00	1.00	0.000	1.00	0.000	

FIGURA 4.2- ARL en función de la magnitud de desviación constante. Tamaño de muestra  $n=5$ .



Con respecto al gráfico CUSUM, resulta más efectivo que los otros dos procedimientos si la media del proceso sufre una desviación inferior a  $\sigma$ . Sin embargo, para desviaciones superiores, este gráfico presenta una potencia inferior a cualquiera de los otros dos.

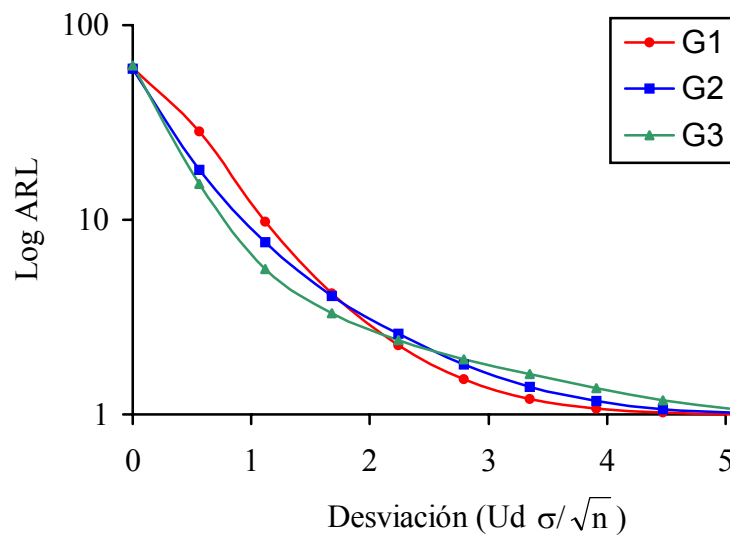
En el rango de desviaciones simuladas, el gráfico de medias con señales adicionales presenta una efectividad intermedia entre el gráfico de medias estándar y el CUSUM. Esto implica, que el incremento de potencia que supone la utilización de señales adicionales cuando la desviación es pequeña va acompañado de una disminución en su efectividad frente a desviaciones grandes. Se pierde así parte de la ventaja que el gráfico de medias de Shewhart tiene respecto al CUSUM, ya que si bien este último se caracteriza por una mayor rapidez en la detección de desviaciones pequeñas, el primero presenta ventajas cuando la desviación es importante, aunque esto último también depende de los parámetros  $k$  y  $h$  elegidos para el CUSUM.

Los resultados obtenidos para otros tamaños de muestra son similares a los anteriores, como en principio era de esperar. El tamaño de muestra determina la magnitud de desviación constante para la que un procedimiento resulta más efectivo que

otro. En la FIGURA 4.3, se recogen valores del ARL de los tres procedimientos gráficos, en función de la magnitud de desviación en la media, expresada en este caso en unidades de desviación típica de la media muestral,  $\sigma/\sqrt{n}$ .

Se observa que la desviación por debajo de la cual la introducción de señales adicionales mejora la efectividad del gráfico de medias es aproximadamente igual a  $1.7\sigma/\sqrt{n}$ , siendo  $n$  el tamaño de muestra considerado. Análogamente, el gráfico CUSUM considerado en nuestro estudio resulta ser más efectivo que los otros dos procedimientos cuando esta desviación es inferior a  $2.5\sigma/\sqrt{n}$ . Por encima de este valor, el gráfico de medias con señales adicionales presenta un comportamiento intermedio entre el CUSUM elegido y el gráfico de medias estándar.

FIGURA 4.3- ARL en función de la magnitud de desviación constante, expresada en unidades de  $\sigma/\sqrt{n}$ .



De los resultados expuestos se desprende que el gráfico de control a adoptar en cada situación depende del orden de magnitud de la desviación constante cuya detección precoz resulta relevante. Por otro lado, se observa que la introducción de señales adicionales en el gráfico de medias mejora su efectividad frente a desviaciones de poca magnitud. Sin embargo, el gráfico CUSUM diseñado con unos parámetros  $k$  y  $h$

adecuados sigue siendo el procedimiento más potente en este caso. Para el rango de desviaciones estudiado, las propiedades del gráfico de medias con señales adicionales resultan intermedias entre lo otros dos gráficos considerados.

Teniendo en cuenta lo anterior, si la magnitud de desviación crítica es grande no queda justificada estadísticamente la introducción en el gráfico de medias de señales adicionales de falta de control. En este caso, la utilización de estas señales supone una mayor complejidad de manejo, sin que mejore la efectividad del gráfico para la situación de interés.

#### **4.4.2- Desviación progresiva en la media**

Otro tipo de salida de control que puede darse en el proceso es una desviación en la media que en lugar de provocar un cambio o "salto" a otro nivel, se manifieste de forma que su valor varía con el tiempo. Este podría ser el caso, por ejemplo, de un proceso de fabricación en el que tiene lugar desgaste de herramienta, de un reactivo, etc...

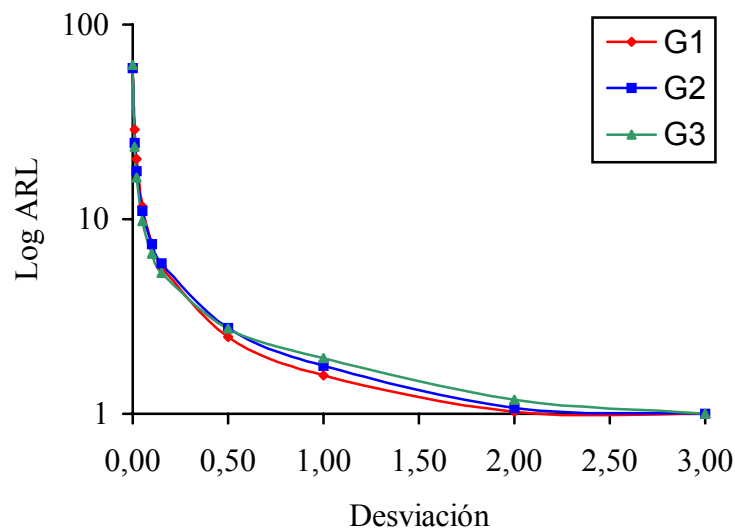
Sin pérdida de generalidad, se va a asumir que esta segunda situación consiste en una deriva progresiva y que una vez producida la salida de control la media del proceso es igual a  $m_0 + \delta\sigma_i$ , siendo  $i$  el número de muestra obtenida desde que empezó el cambio. En este caso, la magnitud de desviación  $\delta\sigma$  expresa la desviación ocurrida en la media por intervalo de muestreo y se supondrá también constante.

Mediante el procedimiento de simulación descrito, se han estudiado las propiedades de los gráficos de medias estándar, con señales adicionales y CUSUM, asumiendo este tipo de salida de control. Los resultados correspondientes se reflejan en la TABLA 4.2 y FIGURA 4.4, para tamaño de muestra  $n=5$ . Se observa en la tabla y el gráfico mencionados que cuando la desviación en la media del proceso es progresiva, y el tamaño de muestra es  $n=5$ , las diferencias entre los tres procedimientos son mucho menores que en el caso de desviación constante.

TABLA 4.2- ARL en función de la magnitud de desviación progresiva. Tamaño de muestra n=5.

$\delta$ (Ud $\sigma$ )	G1		G2		G3	
	ARL	S <sub>ARL</sub>	ARL	S <sub>ARL</sub>	ARL	S <sub>ARL</sub>
0.00	60.00	0.000	59.89	0.538	61.97	0.587
0.01	28.84	0.164	24.57	0.135	23.56	0.108
0.02	20.38	0.099	17.70	0.084	16.30	0.063
0.05	11.65	0.048	11.03	0.043	9.80	0.032
0.10	7.51	0.027	7.46	0.025	6.62	0.018
0.15	5.71	0.019	5.91	0.018	5.28	0.013
0.50	2.49	0.008	2.75	0.008	2.73	0.006
1.00	1.57	0.005	1.76	0.005	1.93	0.004
2.00	1.02	0.001	1.07	0.005	1.18	0.003
3.00	1.00	0.000	1.00	0.000	1.00	0.000

FIGURA 4.4 - ARL en función de la magnitud de desviación progresiva, expresada en unidades de desviación típica del proceso. Tamaño de muestra n=5.



Para valores de la deriva inferiores a  $0.15\sigma$ , ó en general a  $0.335 \sigma/\sqrt{n}$ , el procedimiento óptimo, entendiendo como tal el que minimiza el valor del ARL, es el

gráfico CUSUM. En este mismo rango de desviaciones, el gráfico de medias con señales adicionales presenta una efectividad intermedia entre el procedimiento estándar y el CUSUM.

Para desviaciones progresivas de mayor magnitud, el gráfico de medias estándar resulta ser el procedimiento más efectivo. Como ya hemos señalado, sin embargo, las diferencias son en todo caso de poca magnitud.

Teniendo en cuenta lo anterior, la elección del procedimiento de control cuando la desviación crítica en la media del proceso es progresiva, debe basarse en criterios distintos a los estadísticos, puesto que en este sentido las propiedades de los tres gráficos analizados son muy similares.

#### **4.5- FUNCIÓN DE PESO DE LOS PROCEDIMIENTOS COMPARADOS**

Según se ha visto en el apartado anterior, las propiedades estadísticas del gráfico de medias con señales adicionales, expresadas en función del ARL bajo distintas hipótesis de estado en el proceso, resultan intermedias entre las del gráfico de medias estándar y el CUSUM. La justificación de este resultado puede encontrarse en la manera con que cada gráfico utiliza las observaciones.

En el gráfico de medias propuesto por Shewhart la señal de falta de control viene dada por la aparición de un punto fuera de los límites de control. Esto quiere decir que el peso de la decisión recae únicamente en el último punto graficado.

Para el gráfico CUSUM, la señal de falta de control depende del valor de la suma acumulada de todas las diferencias entre las observaciones y la media, sumando o restando a las mismas el valor de referencia según se pretenda detectar desviaciones positivas o negativas de la media del proceso. Por lo tanto, las observaciones recientes tienen el mismo peso que las obtenidas en instantes de tiempo alejados.

Sin embargo, cuando se utilizan señales adicionales de falta de control la influencia de las observaciones pasadas varía según hayan sido definidas. Así, para el grupo de señales adicionales que nos ocupa, la Señal 1 considera sólo el último punto graficado, la Señal 2 toma en cuenta los tres últimos puntos y la 3 y 4 los siete últimos. La TABLA 4.3 recoge los pesos aproximados que cada una de estas señales dan a observaciones pasadas, siendo  $t$  el instante actual, asumiendo en cada caso que sólo se utiliza una señal. Por ejemplo, los valores de la Señal 4 están condicionados a que no se haya utilizado ninguna de las otras tres señales.

TABLA 4.3- Peso que con cada señal de falta de control se da a las observaciones en el gráfico.

TIEMPO	$t$	$t-1$	$t-2$	$t-3$	$t-4$	$t-5$	$t-6$	$t-7$
Señal 1	1	0	0	0	0	0	0	0
Señal 2	1/3	1/3	1/3	0	0	0	0	0
Señal 3	1/7	1/7	1/7	1/7	1/7	1/7	1/7	0
Señal 4	1/7	1/7	1/7	1/7	1/7	1/7	1/7	0

Para el cálculo del peso correspondiente a cada observación cuando se utilizan conjuntamente las señales anteriores, vamos a asumir la simplificación consistente en sumar los valores reflejados en la tabla anterior para cada instante. Evidentemente esto no da el valor real, pero si una idea aproximada de la operación del gráfico cuando se combinan estas señales. Los valores resultantes de recogen en la TABLA 4.4 y en la correspondiente FIGURA 4.5.

Se observa en la tabla y figura mencionadas, que el gráfico de medias con señales de falta de control adicionales es equivalente a un procedimiento que considera la información contenida en observaciones pasadas, pero dándoles menor peso según se alejan en el tiempo. En este sentido, su comportamiento es parecido al del gráfico EWMA, que se estudian en el capítulo 5 de esta tesis.

Resulta, por tanto, una situación intermedia respecto al gráfico de medias estándar, que sólo considera el último punto graficado, y el gráfico CUSUM, que da el mismo peso a todas las observaciones pasadas. Así, sus propiedades estadísticas son también intermedias entre las de estos dos procedimientos.

TABLA 4.3- Pesos totales por observación cuando se combinan las cuatro señales. La última columna recoge los valores de la segunda divididos por la suma total de pesos, que en este caso es igual a 4.

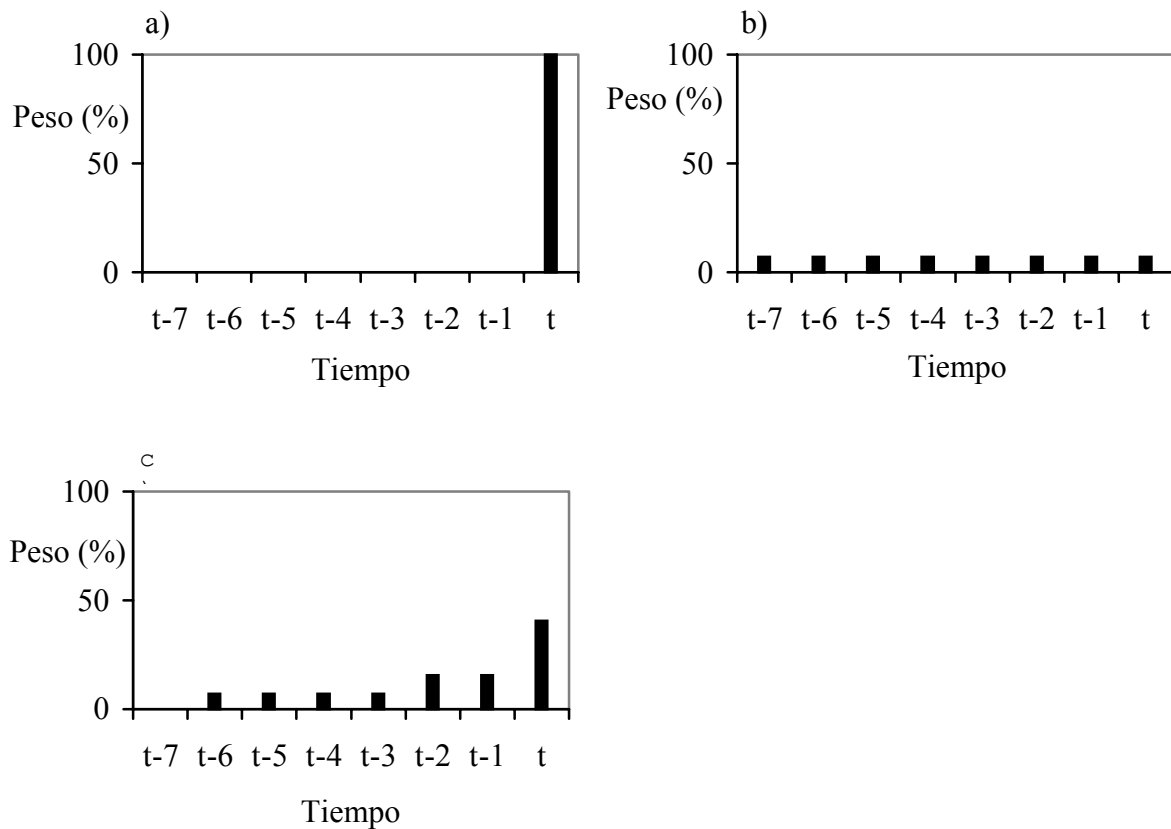
TIEMPO	PESO TOTAL	PESO EST.
t	34/21	0.4047
t-1	13/21	0.1548
t-2	13/21	0.1548
t-3	2/7	0.0714
t-4	2/7	0.0714
t-5	2/7	0.0714
t-6	2/7	0.0714

#### 4.6- FRECUENCIA RELATIVA DE LAS SEÑALES DE FALTA DE CONTROL

Del análisis realizado en apartados anteriores se desprende que cuando se incluyen señales adicionales las propiedades estadísticas del gráfico de medias se ven muy modificadas. Así, este procedimiento resulta más efectivo frente a desviaciones constantes de poca magnitud en la media del proceso, suponiendo también este aumento de sensibilidad una mayor cantidad de falsas alarmas y una menor potencia cuando las desviaciones son grandes.



FIGURA 4.5- Función de peso de a) Gráfico de medias estándar, b) Gráfico CUSUM y c) Gráfico de medias con señales adicionales.



Una cuestión importante que puede ayudar a comprender mejor este cambio de propiedades, es determinar cómo contribuye cada una de las señales al valor del ARL en cada caso. Por otro lado, si en el análisis de esta cuestión se observa que alguna de las señales se presenta con muy poca frecuencia para el rango de desviaciones pequeñas, podrá simplificarse el gráfico, eliminándola en la práctica.

#### 4.6.1- Desviación constante en la media

Con el fin de cuantificar estas contribuciones, se ha registrado en cada simulación el número de veces que la salida de control se debió a cada señal. A partir de este valor se ha calculado su frecuencia relativa de aparición para cada desviación en la media. Los resultados para tamaño de muestra  $n=5$  se recogen en la TABLA 4.5 y en la

correspondiente FIGURA 4.6. Las frecuencias relativas se han expresado como tanto por cien de veces que se presenta cada señal respecto al total de las señales.

TABLA 4.5- Frecuencia relativa de las señales adicionales de falta de control en función de la magnitud de desviación constante de la media. Tamaño de muestra  $n=5$ . Entre paréntesis la desviación típica de la estimación.

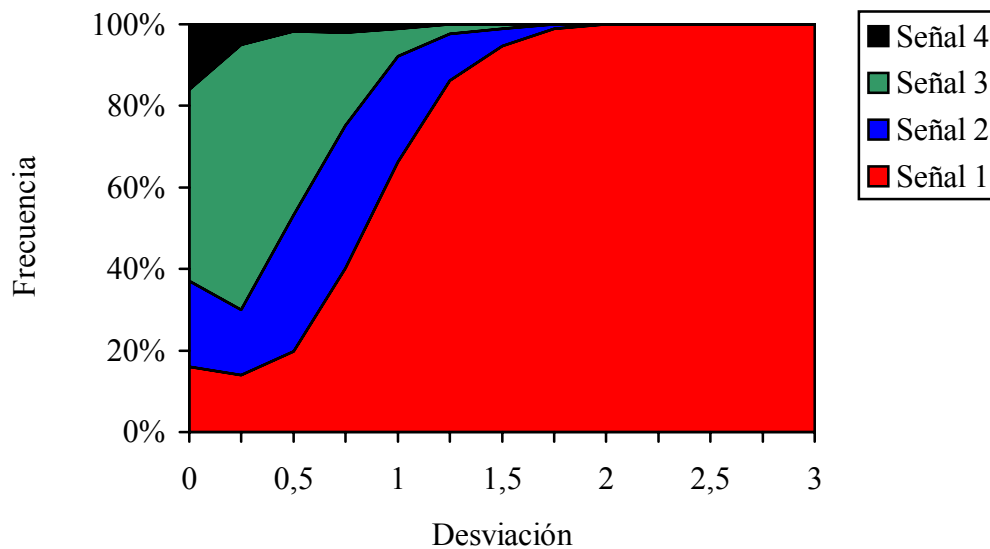
$\delta$ (Ud $\sigma$ )	Señal 1	Señal 2	Señal 3	Señal 4
0.00	16 (0.3666)	21 (0.4073)	47 (0.4990)	16 (0.3666)
0.25	14 (0.3470)	16 (0.3666)	65 (0.4769)	5 (0.2179)
0.50	22 (0.4142)	26 (0.4386)	50 (0.5000)	2 (0.1400)
0.75	39 (0.4877)	37 (0.4828)	22 (0.4142)	2 (0.1400)
1.00	59 (0.4918)	34 (0.4737)	6 (0.2375)	1 (0.0995)
1.25	75 (0.4330)	23 (0.1771)	2 (0.1400)	0
1.50	89 (0.3129)	10 (0.3000)	1 (0.0995)	0
1.75	96 (0.1959)	4 (0.1959)	0	0
2.00	99 (0.0995)	1 (0.0995)	0	0
2.25	100	0	0	0
2.50	100	0	0	0
2.75	100	0	0	0
3.00	100	0	0	0

En la tabla y figura mencionadas se observa que la frecuencia relativa de aparición de cada señal depende de la magnitud de desviación constante de la media del

proceso. Cuando esta desviación es pequeña, la mayor parte de las señales de falta de control se deben a la aparición de rachas en el gráfico. Así, la Señal 3 presenta una frecuencia relativa alta y en segundo lugar la Señal 2.

En concreto, para un tamaño de muestra  $n=5$  las señales de racha son más importantes que la Señal 1, punto fuera de los límites de control, para desviaciones aproximadamente inferiores a  $0.75\sigma$ . Para cualquier otro  $n$ , esto último sucede cuando la desviación es inferior a  $1.67\sigma/\sqrt{n}$ . Por tanto, cuando la magnitud de desviación en la media del proceso supera este valor, la mayor parte de las señales de falta de control se deben a la Señal 1.

FIGURA 4.6- Frecuencia relativa de cada señal. Tamaño de muestra  $n=5$ .



Éste primer resultado es coherente con lo señalado en el apartado anterior al comparar la efectividad del gráfico de medias estándar y la del gráfico de medias con señales adicionales. Según se vio, las señales adicionales incrementan la potencia del gráfico para desviaciones inferiores a  $1.7\sigma/\sqrt{n}$ , presentando para desviaciones superiores a este valor un ARL más bajo el gráfico de medias estándar. Para ajustar el ARL bajo control a 60, los límites en el gráfico de medias estándar se han tomado a una

distancia  $2.39 \sigma/\sqrt{n}$  de la media del proceso, mientras que en el gráfico de medias con señales adicionales estos límites están situados a una distancia  $3 \sigma/\sqrt{n}$ . Por tanto, a desviaciones altas el ARL del gráfico estándar será menor por ser mayor la probabilidad de que tenga lugar la aparición de un punto fuera de los límites.

Otro resultado importante es la evolución que sigue la frecuencia relativa de aparición de la Señal 4, racha ascendente o descendente de 7 puntos seguidos. Según se observa, la frecuencia más alta se da cuando el proceso está bajo control, siendo del mismo orden de magnitud que la frecuencia relativa correspondiente a la Señal 1, aparición de un punto fuera de los límites. Esta señal se presenta también en el caso de desviaciones de muy poca magnitud, y con una frecuencia prácticamente cero. Por tanto, contribuye al incremento de falsas alarmas pero muy poco al aumento de efectividad registrado en el gráfico.

Con el fin de comprobar esta última afirmación, se ha repetido la simulación del gráfico de medias pero incluyendo sólo las señales 1, 2 y 3. Los resultados se recogen en la TABLA 4.6.

Se observa en dicha tabla que al eliminar la Señal 4 el ARL bajo control aumenta significativamente, lo que implica en la práctica una menor cantidad de falsas alarmas. Por otro lado, los ARL del procedimiento resultante de incluir las tres señales adicionales, son a desviaciones bajas ( $< 0.75\sigma$  ó  $< 1.67 \sigma/\sqrt{n}$ ) algo mayores que los correspondientes al mismo gráfico con las cuatro señales, aunque las diferencias son poco relevantes. Sin embargo, a desviaciones altas, tal y como era de esperar el comportamiento de los dos procedimientos es muy similar.

Por lo tanto, la inclusión de la Señal 4 no está justificada desde el punto de vista estadístico, ya que su contribución a la mejora de efectividad del gráfico es muy pequeña y redundante en una cantidad de falsas alarmas algo mayor que cuando no se considera.

Al eliminar esta señal, la frecuencia relativa de la Señal 3, racha de siete puntos por encima o por debajo de la media, aumenta según se observa al comparar los valores de la TABLA 4.5 con los de la TABLA 4.7.

TABLA 4.6- ARL del gráfico de medias con 3 y 4 señales de falta de control, en función de la magnitud de desviación constante en la media. Tamaño de muestra  $n=5$ .

$\delta$ (Ud $\sigma$ )	4 señales		3 señales	
	ARL	$S_{ARL}$	ARL	$S_{ARL}$
0.00	59.89	0.558	64.92	0.648
0.25	18.05	0.157	19.60	0.178
0.50	7.65	0.053	8.37	0.059
0.75	4.07	0.025	4.28	0.026
1.00	2.59	0.015	2.57	0.015
1.25	1.81	0.009	1.81	0.009
1.50	1.38	0.006	1.40	0.006
1.75	1.17	0.004	1.18	0.004
2.00	1.06	0.002	1.06	0.002
2.25	1.02	0.001	1.02	0.001
2.50	1.00	0.000	1.00	0.000
2.75	1.00	0.000	1.00	0.000
3.00	1.00	0.000	1.00	0.000

#### 4.6.2- Deriva progresiva en la media

Uno de los resultados del análisis anterior es que cuando la desviación en la media es constante no está justificada la inclusión de la Señal 4. Sin embargo, en algunas situaciones la salida de control del proceso puede deberse no ya a un cambio de nivel en la media, sino a la aparición de una deriva progresiva en la misma. Según se vio en el apartado anterior, en este caso el incremento de potencia debido a las señales adicionales es mucho menor, aunque se sigue manifestando para derivas de poca magnitud.

Teniendo en cuenta la naturaleza de esta salida de control, se podría pensar que la contribución de la Señal 4 será mayor con deriva progresiva. Con el fin de estudiar este punto, se ha obtenido en las simulaciones con deriva progresiva la frecuencia relativa de cada señal, asumiendo distintas magnitudes para la desviación. Los resultados se reflejan en la TABLA 4.8, para el caso de tamaño de muestra  $n=5$ .

TABLA 4.7- Frecuencia relativa de aparición, expresada en porcentaje, de las tres señales de falta de control en función de la magnitud de desviación constante. Tamaño de muestra  $n=5$ .

$\delta$ (Ud $\sigma$ )	Señal 1	Señal 2	Señal 3
0.00	15 (0.3571)	20 (0.4000)	65 (0.4769)
0.25	14 (0.3470)	17 (0.3756)	69 (0.4625)
0.50	21 (0.4073)	26 (0.4386)	53 (0.4991)
0.75	38 (0.4854)	37 (0.4828)	25 (0.4330)
1.00	59 (0.4918)	34 (0.4737)	7 (0.2551)
1.25	75 (0.4330)	23 (0.1771)	2 (0.1400)
1.50	89 (0.3129)	10 (0.3000)	1 (0.0995)
1.75	96 (0.1959)	4 (0.1959)	0
2.00	99 (0.0995)	1 (0.0995)	0
2.25	100	0	0
2.50	100	0	0

Se observa en la tabla mencionada que, cuando se produce en la media una desviación progresiva de poca magnitud, la frecuencia relativa de las distintas señales de falta de control es muy similar al caso de desviación constante. Es decir, la mayor parte de las señales de falta de control se deben en este caso a la aparición de rachas (señales 2, 3), presentándose un menor número de veces el punto fuera de los límites. Sin embargo, para derivas  $>0.15\sigma$  ( $0.335 \sigma/\sqrt{n}$ ) esta última señal es la más frecuente, resultado que coincide con lo mencionado en el apartado anterior, ya que para esta misma desviación la efectividad del gráfico de medias con señales adicionales es menor que la del gráfico de medias estándar.

TABLA 4.8- Frecuencia relativa de las señales de falta de control, en función de la magnitud de desviación progresiva. Tamaño de muestra  $n=5$ .

$\delta$ (Ud $\sigma$ )	Señal 1	Señal 2	Señal 3	Señal 4
0.01	14	16	62	8
0.02	16	20	58	6
0.05	25	26	45	4
0.10	35	32	30	3
0.15	47	33	18	2
0.50	80	17	3	0
1.00	93	6	1	0
2.00	99	1	0	0
3.00	100	0	0	0

Por otra parte, y de forma similar a como ocurría con desviación constante, la Señal 4 se presenta con una frecuencia relativa baja, aunque ligeramente superior cuando la deriva es pequeña. Esto implica que tampoco en este caso está justificada su inclusión, ya que la mejora en efectividad que se deriva de la misma es poco relevante.

Con respecto a los otros tamaños de muestra, se observa que existe un paralelismo entre la evolución de las frecuencias relativas a distintas desviaciones para

un tamaño de muestra  $n$  dado y la evolución de dichas frecuencias tomando una desviación determinada  $\delta$  y tamaños de muestra crecientes. Este paralelismo era de esperar dado que un incremento de  $n$  para un valor de  $\delta$  fijo es equivalente a un incremento de la desviación considerada en relación a la desviación típica de la media muestral. Así, se constata que para una desviación dada, la Señal 3 y la Señal 2 aparecen con bastante frecuencia para tamaños de muestra  $n$  pequeños. A medida que aumenta el valor de  $n$ , la Señal 1 resulta ser la más frecuente, tal y como ocurre para el caso de desviaciones grandes.

#### **4.6.3- Frecuencia de las señales de falta de control cuando aumenta la dispersión del proceso**

Las frecuencias relativas de las señales de falta de control se han obtenido asumiendo un cambio en la media, puesto que el gráfico de medias se utiliza para la detección de causas especiales de variabilidad que puedan afectar a la posición del proceso.

Normalmente, el gráfico de medias se construye y maneja en paralelo con el gráfico de recorridos, para control de la dispersión. En principio, es de esperar que si la salida de control en el proceso se debe a un cambio de la dispersión, aparezcan rachas o pautas anómalas en el gráfico de recorridos.

Sin embargo, puede ocurrir que este tipo de salida se manifieste también en el gráfico de medias, mediante la aparición de alguna de las señales hasta aquí consideradas. Con el fin de cuantificar este último punto, se han realizado algunas simulaciones, en las que la salida de control se debe a un cambio en la dispersión en lugar de en la media.

En cada una de estas simulaciones se ha registrado la frecuencia de aparición de cada señal, obteniéndose los resultados que se reflejan en la TABLA 4.9. Las frecuencias están expresadas como porcentajes;  $\sigma_r$  y  $\sigma_a$  representan respectivamente los valores de



la desviación típica antes y después de producirse la desviación. Se ha asumido que en cada salida de control aumenta la dispersión.

TABLA 4.9- Frecuencia relativa de las señales de falta de control en función del aumento relativo de la dispersión. Tamaño de muestra  $n=5$ . Entre paréntesis la desviación típica de la proporción estimada.

$\sigma_r/\sigma_0$	Señal 1	Señal 2	Señal 3	Señal 4
1.02	15.7 (0.3638)	18.8 (0.3907)	49.7 (0.4999)	15.8 (0.3647)
1.04	16.2 (0.3684)	20.1 (0.4007)	49.4 (0.4999)	14.3 (0.3501)
1.06	21.3 (0.4094)	23.1 (0.4215)	42.0 (0.4935)	13.6 (0.3428)
1.08	24.2 (0.4283)	26.5 (0.4413)	36.3 (0.4809)	13.0 (0.3363)
1.1	24.3 (0.4289)	26.8 (0.4429)	37.3 (0.4836)	11.6 (0.3190)
1.2	34.3 (0.4747)	33.4 (0.4716)	24.5 (0.4301)	7.8 (0.2682)
1.3	41.2 (0.4922)	33.2 (0.4709)	19.9 (0.3992)	5.7 (0.2318)
1.4	48.3 (0.4997)	34.2 (0.4744)	13.4 (0.3406)	4.1 (0.1983)
1.5	50.1 (0.5000)	32.4 (0.4680)	14.0 (0.3468)	3.5 (0.1838)
1.6	57.0 (0.4951)	31.3 (0.4637)	9.3 (0.2904)	2.4 (0.1530)
2.0	69.5 (0.4604)	24.1 (0.4277)	5.2 (0.0493)	1.2 (0.1088)
2.5	78.2 (0.4129)	17.4 (0.3791)	3.7 (0.1888)	0.7 (0.0834)

Se observa que mientras la dispersión del proceso no aumente más de un 10% del valor inicial, la apariencia del gráfico de medias con señales adicionales será aproximadamente la misma que estando bajo control para la media. Esto se manifiesta

con una frecuencia de aparición de rachas relativamente alta. Para desviaciones mayores, como era de esperar, aumenta el número de puntos fuera de los límites de control.

Destaca, por otro lado, la tendencia que sigue la frecuencia de aparición de la Señal 2, que sufre una disminución mucho más gradual que en el caso de desviaciones en la media. Así, cuando se produce un aumento del 50% aproximadamente en el valor de la dispersión, 1/3 de las señales que se producen en el gráfico de medias se deben a esta señal.

La explicación puede encontrarse en su misma definición, pues en esta señal se considera falta de control cuando aparecen dos de tres puntos seguidos entre los límites de aviso y los de control, pudiendo estar los dos puntos a distintos lados de la línea central. Es lógico, por tanto, que aumente su frecuencia conforme lo hace la dispersión del proceso.

De estos últimos resultados se concluye que la aparición de rachas en el gráfico de medias puede deberse también a cambios en la dispersión del proceso. Por tanto, con el fin de evitar el sobreajuste excesivo, se debe contrastar la apariencia del gráfico de medias con la del gráfico de recorridos, para investigar si existe alguna causa especial de variabilidad y si esta afecta a la media o a la dispersión del proceso.

#### **4.7 - OTRAS SEÑALES DE FALTA DE CONTROL**

Una conclusión a extraer de todo lo anterior es que la introducción de señales adicionales en el gráfico de medias mejora su efectividad frente a desviaciones constantes de poca magnitud. Sin embargo, en este mismo rango de desviaciones sigue siendo el gráfico CUSUM el procedimiento más potente. Por otro lado, la utilización del conjunto de señales adicionales considerado en este análisis presenta como desventaja un aumento considerable de la cantidad de falsas alarmas en el gráfico de medias.

Una alternativa, en aquellas situaciones en las que se prefiere utilizar el gráfico de medias en vez del CUSUM y la magnitud de desviación crítica de la media sea de poca magnitud, es incluir una combinación distinta de señales adicionales de falta de control, diseñada con el fin de no tener una disminución excesivamente grande del ARL bajo control, pero que resulten en un incremento de sensibilidad del gráfico de medias.

Tal y como se señaló en el apartado de revisión bibliográfica, la combinación de señales adicionales más recomendada por la Western Electric Company es utilizada frecuentemente por algunas empresas americanas, habiendo sido estudiadas sus propiedades por diversos autores. Por tanto, tiene sentido plantearse la comparación de este conjunto de señales y el considerado en nuestro análisis con el fin de determinar si el primero presenta alguna ventaja respecto al último.

Esta segunda combinación de señales es la siguiente:

**Señal 1'** : Punto fuera de los límites de control.

**Señal 2'** : Dos de tres puntos seguidos entre los límites de aviso y los de control, pero a un mismo lado de la línea central.

**Señal 3'** : Cuatro de cinco puntos seguidos por encima de  $m_0 \pm \sigma / \sqrt{n}$ .

**Señal 4'** : Una racha de ocho puntos seguidos a un mismo lado de la línea central.

Una diferencia entre los dos grupos de señales estriba en que no existe entre las definidas por la Western Electric Company ninguna parecida a la señal de racha ascendente o descendente de 7 puntos seguidos. También se observa que la Señal 4' exige una racha mínima de 8 puntos, en vez de 7 como sucedía con la Señal 3 de la combinación estudiada, y que la Señal 2' exige que la racha se produzca a un mismo lado de la línea central.

Teniendo en cuenta lo anterior, y los resultados obtenidos respecto a la frecuencia de aparición de rachas, es de esperar que las propiedades del gráfico de medias con este segundo grupo de señales difieran de las correspondientes a este gráfico pero con la combinación analizada en el presente capítulo.

En la TABLA 4.10 se recogen los valores del ARL para el gráfico de medias con la combinación propuesta por la W.E.C. calculados por Champ y Woodall (1987), junto con los ARL del gráfico estándar y del gráfico de medias con las señales adicionales estudiadas en este trabajo, en función de la magnitud de desviación constante en la media expresada en unidades de desviación típica de la media muestral.

En primer lugar, se observa que el ARL bajo control estimado para la combinación de señales de falta de control recomendada por la Western Electric Company es aproximadamente igual a 91. Aunque este resultado supone también una considerable reducción del ARL bajo control, no resulta tan grande como la correspondiente a la combinación de señales estudiadas en la presente tesis.

La justificación de este primer resultado se encuentra, tal y como se señalaba antes, en las diferencias existentes entre los dos grupos de señales. Así, la probabilidad de aparición de la Señal 3 es el doble de la correspondiente a la Señal 4'. También se observa que difieren la Señal 2 y la Señal 2', resultando esta última con una probabilidad más baja. Por tanto, en principio tiene sentido que la cantidad de falsas alarmas cuando se incluye el conjunto de señales propuesto por la Western Electric Company sea algo menor.

Calculando el porcentaje de muestras que resultan en falsa alarma, se obtiene para el gráfico G2 un valor de 1.72%, mientras que para G3 resulta igual a 1.09%, por lo que la diferencia en cuanto a número de sobreajustes entre los dos procedimientos no es excesiva en términos absolutos.

Aunque debido a sus diferentes ARL bajo control, los dos procedimientos no son estrictamente comparables, se va a considerar el comportamiento de cada uno en el rango de desviaciones simuladas.

En general, el gráfico de medias con las señales adicionales analizadas en nuestra investigación presenta una efectividad mayor que el mismo gráfico con las

señales propuestas por la Western Electric Company, siendo esta diferencia más grande para desviaciones bajas. Así, para desviaciones del orden de  $0.4\sigma/\sqrt{n}$ , la diferencia de ARL es igual a 12.57, a favor del primer grupo de señales.

TABLA 4.10- Comparación de la efectividad del gráfico de medias con distintas señales de falta de control. Desviación constante en la media expresada en unidades de desviación típica de la media muestral,  $\sigma/\sqrt{n}$ .

G1: Gráfico de medias estándar con límites a  $m_0 \pm 2.54\sigma/\sqrt{n}$ .

G2: Gráfico de medias con el primer grupo de señales de falta de control.

G3: Gráfico de medias con las señales recomendadas por la W.E.C.

Entre paréntesis la desviación típica del ARL estimado para G2.

$\delta$ (Ud $\sigma/\sqrt{n}$ )	GRÁFICOS DE CONTROL		
	G1	G2	G3
0.0	92.14	58.06 (0.575)	91.75
0.2	78.66	35.05 (0.334)	66.80
0.4	56.12	24.04 (0.219)	36.61
0.6	36.99	16.28 (0.141)	20.90
0.8	24.18	11.36 (0.092)	13.25
1.0	16.13	9.23 (0.069)	9.22
1.2	11.09	6.43 (0.044)	6.89
1.4	7.86	5.30 (0.033)	5.41
1.6	5.76	4.45 (0.033)	4.41
1.8	4.35	3.72 (0.024)	3.68
2.0	3.39	3.09 (0.019)	3.13
2.2	2.72	2.63 (0.031)	2.70
2.4	2.25	2.28 (0.026)	2.35
2.6	1.91	2.02 (0.023)	2.07
2.8	1.66	1.80 (0.020)	1.85
3.0	1.48	1.62 (0.008)	1.67

Adicionalmente, Champ y Woodall (1987) concluyen en su análisis que el gráfico de medias con este grupo de señales es menos efectivo que el CUSUM para un rango de desviaciones pequeñas.

Teniendo en cuenta lo anterior, se puede considerar que la combinación propuesta por la W.E.C. no incrementa la potencia del gráfico de medias tanto como la analizada en apartados anteriores, siendo la frecuencia de falsas alarmas de un orden de magnitud similar.

## **CAPÍTULO 5**

### **GRÁFICOS DE CONTROL PARA OBSERVACIONES INDIVIDUALES**

#### **5.1- INTRODUCCIÓN**

Se presentan a menudo situaciones en las que no tiene sentido plantearse la toma de muestras constituidas por varias unidades consecutivas, porque se dispone sólo de una serie de observaciones individuales. Este es el caso de las industrias de procesos, como son, por ejemplo, las del sector químico. En estas industrias, la disponibilidad de una sola medida por intervalo de tiempo puede presentarse en el control de una característica de calidad asociada a un producto fabricado en batch o en continuo, o bien en la medición de parámetros de proceso como temperaturas, ph, etc...

La misma circunstancia se presenta también cuando se controlan magnitudes que, por su naturaleza, aparecen en forma de datos acumulados en el tiempo, siendo lo más frecuente disponer de una serie de observaciones individuales. Un ejemplo de este último caso sería el control de consumo de energía, materia prima etc...

Una herramienta utilizada frecuentemente en estos casos para el control on-line de procesos es el gráfico de medidas individuales, que no es más que un caso particular del gráfico de medias estándar con tamaño de muestra  $n=1$ . Como alternativa a este procedimiento, se emplea, entre otros, el gráfico de medias móviles, que es prácticamente el procedimiento más aplicado en el contexto de observaciones individuales.

Si se opta por este último tipo de gráfico, se debe elegir en la etapa de diseño el tamaño de agrupamiento  $n$  con que se van a calcular las medias a registrar. En este caso, el criterio que permite elegir el tamaño de agrupamiento óptimo debe basarse en las

propiedades estadísticas del gráfico resultante, ya que sea cual sea el valor de  $n$ , todos los procedimientos son equivalentes en cuanto a coste de muestreo, puesto que parten de una misma serie de observaciones individuales.

Aunque el gráfico de medias móviles fue introducido hace tiempo (Roberts, 1958), se desprende de la revisión bibliográfica la no existencia de estudios que aborden el análisis de sus propiedades estadísticas cuando varía el tamaño  $n$  de agrupamiento. Esto implica, entre otras cosas, la carencia de un criterio estadístico que permita elegir el  $n$  apropiado en cada situación.

Teniendo en cuenta lo anterior, uno de los objetivos a cubrir en la tercera parte de nuestra investigación es la determinación del comportamiento y propiedades estadísticas del gráfico de medias móviles, para distintos valores de  $n$  y desviaciones constantes en la media del proceso. Se asume, para la característica de calidad, una distribución aproximadamente normal y que las observaciones son independientes entre sí.

El parámetro a utilizar como medida de su efectividad es también en este caso el ARL, o número medio de observaciones a tomar desde que se produce la salida de control hasta que se detecta en el gráfico. Cuando el proceso está bajo control, el ARL es una medida aproximada de la cantidad de falsas alarmas que se producen. Para su cálculo en cualquiera de los dos casos, se ha recurrido a la simulación en ordenador, según el procedimiento que se expuso en el apartado de metodología, por no existir aproximación analítica a esta cuestión.

Como alternativa a los gráficos de medidas individuales y de medias móviles se recomienda también, cuando se dispone de una sola observación por unidad de tiempo, utilizar el gráfico CUSUM o el EWMA. Estos dos procedimientos se caracterizan por una mayor complejidad de diseño e interpretación, no estando tan extendida su utilización como en el caso de las medias móviles o medidas individuales.



El gráfico CUSUM, aplicable en algunos contextos como herramienta óptima para control de procesos, ha sido objeto de estudio por diversos autores. Sin embargo, el gráfico EWMA, utilizado frecuentemente para realizar predicciones en el campo de la econometría, sólo ha sido considerado como herramienta para control on-line de procesos en algunos artículos recientes, a pesar de que se introdujo con este fin hace tiempo (Roberts, 1959). Recogiendo la recomendación respecto a su uso como herramienta de control en el caso que nos ocupa, es importante estudiar si realmente resultan más efectivos que los procedimientos estándar y si este incremento en potencia compensa la mayor complejidad que supone su utilización.

Teniendo en cuenta lo anterior, una vez analizadas las propiedades estadísticas del gráfico de medias móviles, se aborda un estudio comparativo de todos los gráficos recomendados habitualmente para control de procesos en los que se dispone de una serie de observaciones individuales. Se pretende determinar el procedimiento óptimo en este caso, entendiendo como tal el que minimiza el tiempo que tarda en detectarse una salida de control que se manifiesta mediante un cambio de nivel en la media del proceso.

## **5.2- PROPIEDADES ESTADÍSTICAS DE LOS GRÁFICOS DE MEDIAS MÓVILES**

### **5.2.1- Consideraciones previas**

La construcción y fundamento de este tipo de gráficos se recoge en el Anejo 1. Se desprende de lo expuesto en el mismo, que la primera diferencia importante con respecto al gráfico de medias estándar es que los puntos a registrar, las medias móviles, no son independientes entre sí. Así, la media  $i$  se calcula promediando la observación individual obtenida en el instante  $i$  con  $n-1$  observaciones anteriores. Por lo que el coeficiente de correlación entre dos medias  $i$  e  $i-j$  depende del tamaño de agrupamiento  $n$ , siendo distinto de cero si  $j$  es menor que  $n$ . La expresión del mismo se puede obtener a partir de:

$$\rho(\bar{X}_i, \bar{X}_{i-j}) = \frac{\sigma^2(\bar{X}_i, \bar{X}_{i-j})}{\sigma(\bar{X}_i)\sigma(\bar{X}_{i-j})}$$

Si el tamaño de agrupamiento es  $n$ ,  $\sigma(\bar{X}_i) = \sigma(\bar{X}_{i-j}) = \sigma/\sqrt{n}$ . Por otro lado cuando las observaciones son independientes entre sí, la covarianza entre dos medias móviles es igual a:

$$\sigma^2(\bar{X}_i, \bar{X}_{i-j}) = E[(\bar{X}_i - m)(\bar{X}_{i-j} - m)]$$

donde teniendo en cuenta cómo se calcula cada media,

$$\sigma^2(\bar{X}_i, \bar{X}_{i-j}) = \frac{1}{n^2} E[(X_i + \dots + X_{i-n+1} - nm)(X_i + \dots + X_{i-j-n+1} - nm)]$$

por lo que  $\sigma^2(\bar{X}_i, \bar{X}_{i-j}) = (n-j)\sigma/n^2$ . Sustituyendo este resultado en la expresión del coeficiente de correlación se obtiene:

$$\rho(\bar{X}_i, \bar{X}_{i-j}) = \frac{n-j}{n}$$

Así, el coeficiente de correlación entre dos medias móviles consecutivas  $\bar{X}_i$  y  $\bar{X}_{i-1}$  es igual a  $\rho = 1 - (1/n)$ , incrementándose a medida que  $n$  aumenta.

Este resultado se podría haber obtenido también, teniendo en cuenta que la serie de medias móviles generada a partir de observaciones independientes incorrelacionadas, sigue un modelo ARIMA. En concreto, se corresponde con un MA de orden  $n-1$ , siendo  $n$  el tamaño de agrupamiento; los coeficientes  $\theta_i$  del modelo resultan todos iguales a  $-1$ . Efectivamente si, por ejemplo,  $n=3$  la media móvil del instante  $t$ ,  $\bar{X}_t$  se puede expresar como:

$$\bar{X}_t = m + \frac{a_t}{3} + \frac{a_{t-1}}{3} + \frac{a_{t-2}}{3}$$

considerando que una observación individual es  $X_t = m + a_t$ , donde  $a_t$  es la perturbación aleatoria correspondiente al instante  $t$ . Si se define  $a'_t = a_t/3$  y aplica el operador retardo, se obtiene:

$$\bar{X}_t = m + (1 + B + B^2) a'_t$$

expresión que se corresponde con un MA(2), con  $\theta_1 = \theta_2 = -1$ . Este resultado también

implica que la serie de medias a graficar se encuentra en el caso límite de "casi-invertibilidad" en el que el efecto de observaciones pasadas es constante con el retardo. Por otro lado, aplicando las expresiones generales para la función de autocorrelación de este modelo, se deriva el mismo resultado expuesto anteriormente.

Como consecuencia de la correlación existente entre los puntos graficados, el cálculo de la probabilidad de señal de falta de control, punto fuera de los límites, no puede realizarse aplicando las sencillas expresiones vistas para el gráfico de medias estándar. Adicionalmente, las propiedades del procedimiento resultante diferirán de las del gráfico de medias estándar, a pesar de que se construye y se opera en él de forma similar.

En el gráfico de medias estándar, el ARL bajo control no depende del tamaño de muestra elegido. Sin embargo, en el gráfico de medias móviles, el grado de correlación entre las medias móviles varía según sea  $n$ , por lo que es de esperar que la probabilidad de punto fuera de los límites de control estando bajo control dependa también de este tamaño de agrupamiento.

Con el fin de justificar de alguna manera este último aserto, se va a asumir un modelo más sencillo. Se dispone en cada instante de observaciones  $Y_t$  de una variable aleatoria con sólo dos valores posibles, que codificamos como 0 y 1. El valor de esta variable en el instante  $t$  depende del valor que tenía en el instante  $t-1$ , no influyendo observaciones más alejadas en el tiempo, por lo que se trata de un proceso markoviano. Si la  $P(Y_t=1) = p$  y la  $P(Y_t=0) = q$ , puede construirse la siguiente tabla que refleja las probabilidades conjuntas y marginales de esta variable aleatoria.

	$Y_{t-1}=0$	$Y_{t-1}=1$	
$Y_t=0$	$a$	$q-a$	$q$
$Y_t=1$	$q-a$	$p-q+a$	$p$
	$q$	$p$	

Si las observaciones son independientes entre sí, se tiene que:

$$P(Y_{t-1}=0, Y_t=0) = q^2$$

por lo que el valor de  $a$  es en este caso  $a=q^2$ .

En cualquier otra situación, la covarianza entre  $Y_t$  y  $Y_{t-1}$  se calcula como:

$$\begin{aligned} \sigma^2(Y_t, Y_{t-1}) &= E[(Y_t - p)(Y_{t-1} - p)] = E[Y_t Y_{t-1}] + p^2 - E[p Y_{t-1}] - E[p Y_t] = E[Y_{t-1} Y_t] - p^2 = \\ &= a - q^2 \end{aligned}$$

por lo que el coeficiente de correlación resulta ser:

$$\rho(Y_t, Y_{t-1}) = \frac{a - q^2}{pq}$$

Sea  $Z$  otra variable aleatoria definida como el número de observaciones a tomar de la variable  $Y$  hasta que ésta se presenta con valor 1. La esperanza matemática de  $Z$  es igual a:

$$E(Z) = \sum_{i=1}^{\infty} z_i P(Z=z_i)$$

que puede desarrollarse de la forma:

$$\begin{aligned} E(Z) &= P(Z=1) + 2 P(Z=2) + 3 P(Z=3) + \dots = P(Y_1=1) + 2 P(Y_1=0, Y_2=1) + 3 P(Y_1=0, \\ &Y_2=0, Y_3=1) + \dots = p + 2(q-a) + 3 \left(\frac{a}{q}\right)(q-a) + 4 \left(\frac{a^2}{q^2}\right)(q-a) + \dots = p + (q-a)[2 + 3 \left(\frac{a}{q}\right) \\ &+ \dots] \end{aligned}$$

y restándole  $(a/q)E(Z)$  se obtiene:

$$(1-a/q)E(Z) = p(1-a/q) + (q-a)[2 + (a/q) + (a/q)^2 + (a/q)^3 + \dots]$$

Como  $a \leq q$ ,

$$(1-a/q)E(Z) = p(1-a/q) + (q-a)[1 + 1/(1-a/q)]$$

de donde

$$E(Z) = 1 + q^2/(q-a)$$

Despejando  $a$  de la expresión del coeficiente de correlación, el valor esperado de  $Z$  puede calcularse también como:

$$E(Z) = \frac{1 - \rho p}{p(1 - \rho)}$$

En el modelo anterior, el valor 1 puede representar en un instante  $t$  que un punto registrado en el gráfico resulta en una señal de falta de control, y por lo tanto el valor 0 que no se produce. Si los límites se han situado en la posición estándar,  $m_0 \pm 3\sigma/\sqrt{n}$ , se tiene que estando bajo control  $p=0.0027$ , con lo que  $q=0.9973$ . Asimismo,  $E(Z)$  sería el equivalente al ARL bajo control para el gráfico de medias estándar.

Si las observaciones son independientes, lo que se traduce en un coeficiente de correlación igual a 0, sustituyendo estos valores en la última ecuación, resulta  $ARL=E(Z)=369.37$ . Sin embargo, si dos observaciones sucesivas están correlacionadas, con un  $\rho$  por ejemplo igual a 0.5, el valor de  $E(Z)$  resulta a su vez igual a 737.8. Es decir, cuando los puntos están correlacionados positivamente, el ARL bajo control aumenta, sin embargo cuando la correlación es negativa, disminuye. Por ejemplo, para  $\rho=-0.5$ ,  $E(Z)$  resulta igual a 244.61.

Estos dos últimos resultados son aproximados, puesto que la expresión del ARL bajo control se ha derivado asumiendo un modelo muy simplificado. En cualquier caso, sin embargo, ponen de manifiesto que si los puntos que se registran en el gráfico no son independientes, las propiedades del mismo pueden verse muy modificadas.

Tal y como se ha visto antes, éste sería el caso del gráfico de medias móviles. En este tipo de gráfico, el coeficiente de correlación entre dos medias calculadas con alguna observación en común, es siempre distinto de cero, por lo que es de esperar que el ARL bajo control difiera del correspondiente a un gráfico de medias estándar, y adicionalmente este ARL será distinto según varíe el valor del tamaño de agrupamiento  $n$ .

En el siguiente apartado se estudia mediante simulación la relación existente entre los valores del ARL para el gráfico de medias móviles con el tamaño de agrupamiento y con las desviaciones en la media del proceso.

## 5.2.2- Efecto del tamaño de agrupamiento en el ARL

La TABLA 5.1 refleja los valores del ARL del gráfico de medias móviles en función del tamaño de agrupamiento y de la magnitud de desviación constante en la media del proceso. En este caso, se ha estudiado con mayor detalle el rango de desviaciones pequeñas.

Como primer resultado a destacar, se observa en dicha tabla que el ARL bajo control depende del tamaño de agrupamiento elegido. Tal y como refleja la FIGURA 5.1, este parámetro sigue una pauta creciente a medida que aumenta  $n$ . Así, el ARL bajo control para tamaño de agrupamiento  $n=7$  es casi el doble del correspondiente al gráfico de medidas individuales.

En el apartado anterior se discutió asumiendo un modelo muy simplificado cómo puede verse modificado el valor del ARL cuando existe correlación entre las observaciones graficadas en dos instantes sucesivos. Por otro lado, el aumento del ARL bajo control implica que es más probable que una media móvil caiga dentro de los límites, si la media anterior se encuentra en el mismo estado.

Tal y como se ha indicado antes, el cambio en el ARL bajo control conforme varía el tamaño de agrupamiento tiene importantes consecuencias prácticas. En primer lugar, la cantidad de falsas alarmas en un gráfico de medias móviles, con límites en la posición estándar, es mucho menor que en un gráfico de medidas individuales, disminuyendo esta cantidad a medida que aumenta el tamaño de agrupamiento.

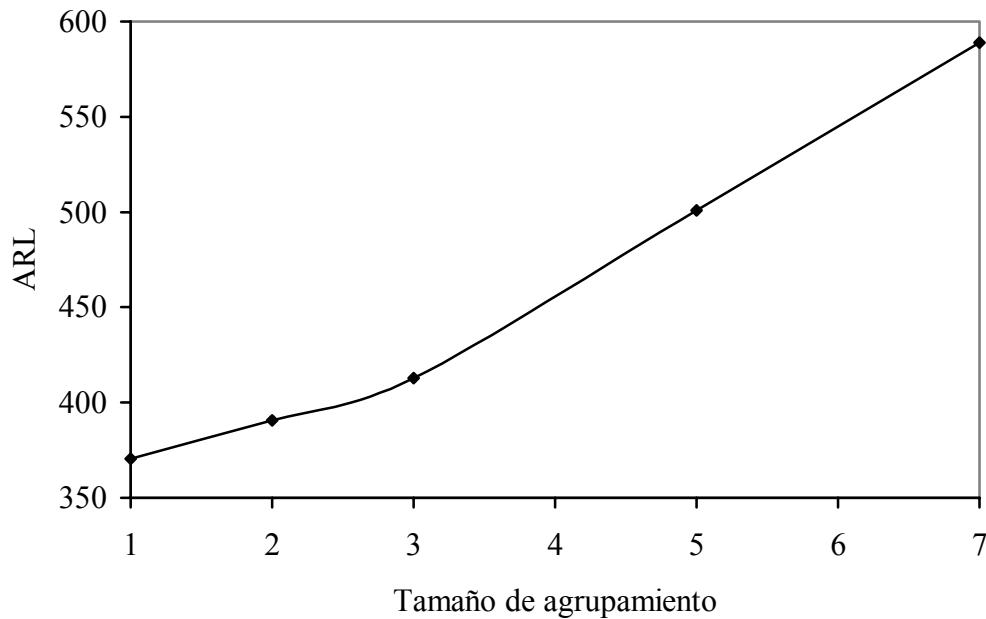
Por otro lado, tal y como se observa en la TABLA 5.1, se produce un aumento del ARL cuando la desviación es pequeña, lo que implica una menor potencia del gráfico frente a este tipo de salidas de control.

Aunque los gráficos no son estrictamente comparables, puesto que presentan un ARL bajo control distinto, se ha construido la FIGURA 5.2, a partir de los resultados de la tabla anterior tomando un tamaño de agrupamiento bajo, como  $n=3$ , y otro alto  $n=7$ .

TABLA 5.1- ARL del gráfico de medias móviles en función del tamaño de agrupamiento n=2, 3, 5 y 7. Entre paréntesis la desviación típica de la estimación

$\delta$ (Ud $\sigma$ )	TAMAÑO DE AGRUPAMIENTO			
	2	3	5	7
0.0	390.77 (3.93)	412.79 (4.160)	500.81 (5.040)	588.87 (5.850)
0.2	265.54 (2.635)	266.6 (2.636)	286.7 (2.918)	296.90 (2.968)
0.3	198.37 (1.979)	189.06 (1.924)	173.75 (1.740)	165.42 (1.643)
0.4	142.57 (1.421)	126.06 (1.233)	108.13 (1.055)	98.33 (0.949)
0.5	102.50 (1.017)	85.41 (0.845)	68.34 (0.667)	60.00 (0.572)
0.6	74.27 (0.733)	60.30 (0.596)	46.68 (0.435)	39.90 (0.366)
0.7	54.78 (0.536)	41.45 (0.404)	31.74 (0.295)	27.51 (0.238)
0.8	40.77 (0.397)	30.65 (0.292)	23.05 (0.206)	20.07 (0.167)
0.9	30.59 (0.301)	22.68 (0.213)	16.94 (0.142)	15.15 (0.116)
1.0	23.05 (0.217)	17.37 (0.157)	13.33 (0.108)	12.23 (0.088)
1.2	14.26 (0.130)	10.99 (0.092)	8.94 (0.063)	8.56 (0.048)
1.6	6.63 (0.056)	5.42 (0.037)	5.14 (0.025)	5.58 (0.020)
1.8	5.68 (0.047)	4.80 (0.031)	4.77 (0.021)	5.25 (0.018)
2.0	3.82 (0.022)	3.57 (0.019)	3.94 (0.013)	4.50 (0.014)
2.2	3.13 (0.019)	3.08 (0.014)	3.58 (0.011)	4.13 (0.012)
2.6	2.34 (0.011)	2.55 (0.009)	3.11 (0.009)	3.60 (0.010)
2.8	2.12 (0.008)	2.39 (0.007)	2.92 (0.008)	3.37 (0.009)
3.0	1.96 (0.007)	2.24 (0.007)	2.75 (0.008)	3.18 (0.009)

FIGURA 5.1- ARL bajo control en función del tamaño de agrupamiento.

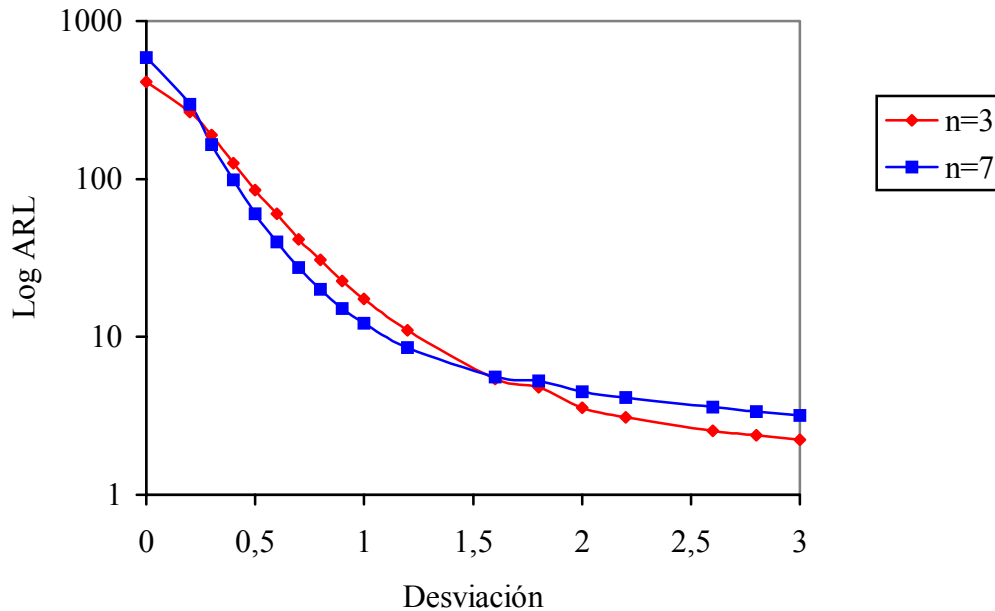


Se observa en la FIGURA 5.2 que para desviaciones muy bajas, en este caso inferiores a  $0.2\sigma$ , resulta más efectivo utilizar un tamaño de agrupamiento pequeño. Cuando la magnitud de desviación es intermedia, es decir entre  $0.3$  y  $1.5\sigma$ , tamaños de agrupamiento grandes implican una mayor rapidez en la detección de la salida de control, lo que se traduce en valores más bajos del ARL. Finalmente, para desviaciones de gran valor el tamaño de agrupamiento óptimo vuelve a ser pequeño.

Este resultado es lógico puesto que cuando se produce la salida de control, si se utiliza un tamaño  $n$  de agrupamiento grande se promedia la primera observación correspondiente a esta situación con un mayor número de observaciones bajo control que en el caso de tamaños de agrupamiento pequeño. Por tanto, en las primeras medias móviles calculadas con  $n$  grande se amortigua más el efecto de la desviación. Así, el valor medio de la media móvil correspondiente a la primera observación después de producirse la salida de control es igual a  $m_0 + (\delta\sigma)/n$ .



FIGURA 5.2- ARL en función de la magnitud de desviación y del tamaño de agrupamiento ( $n=3$  y  $7$ ). Límites de control en la posición estándar  $m_0 \pm 3\sigma/\sqrt{n}$ .



Como consecuencia de esto, cuando la desviación es grande, lo que supone un ARL bajo, con tamaños de agrupamiento pequeños se produce antes la señal de falta de control.

Tal y como se ha señalado antes, los procedimientos resultantes de aplicar distintos tamaños de agrupamiento para el cálculo de las medias móviles, con límites de control en la posición  $m_0 \pm 3\sigma/\sqrt{n}$  no son comparables puesto que sus ARL bajo control difieren.

Con el fin de concretar aún más lo comentado en el párrafo anterior, se ha repetido la simulación ajustando previamente los límites, de forma que los ARL bajo control sean similares al del gráfico de medias estándar, y por lo tanto aproximadamente igual a 370. La TABLA 5.2 recoge la posición L de los límites de control en función del tamaño de agrupamiento elegido.

TABLA 5.2- Posición de los límites para obtener un ARL bajo control aproximadamente igual a 370.

n	1	2	3	4	5	6	7
L	3.00	2.975	2.945	2.907	2.873	2.835	2.807

Teniendo en cuenta los valores reflejados en la tabla anterior, se han repetido las simulaciones del gráfico de medias móviles, tomando distintos tamaños de agrupamiento. La TABLA 5.3 y la correspondiente FIGURA 5.3 recogen algunos de los resultados obtenidos.

En la tabla y figura mencionadas se observa que, una vez ajustados los procedimientos resultantes de aplicar distinto tamaño de agrupamiento, la tendencia seguida por el ARL es similar a la indicada anteriormente. Así, el tamaño de agrupamiento óptimo depende de la magnitud de desviación constante en la media cuya detección precoz es relevante. Cuando esta desviación es pequeña resulta más efectivo agrupar con tamaños relativamente grandes, lo que se justifica por la menor dispersión de la media móvil que la hace más sensible para detectar este cambio de nivel. De los resultados expuestos, se desprende que el orden de magnitud de la desviación para la que esto se cumple, es aproximadamente igual a  $1.25\sigma$ .

Para desviaciones mayores, las diferencias en efectividad entre los procedimientos analizados, son poco relevantes. A pesar de ello, se observa que cuando esta desviación es grande el tamaño de agrupamiento óptimo es pequeño.

La justificación de este resultado se encuentra, tal y como se indicó anteriormente, en que los tamaños de agrupamiento grandes amortiguan la desviación en las primeras medias móviles que se calculan después de producirse la salida de control, puesto que se promedia las observaciones fuera de control con un mayor número de datos anteriores a la desviación.

TABLA 5.3- ARL en función del tamaño de agrupamiento y la magnitud de desviación constante en la media.

$\delta$ (Ud $\sigma$ )	TAMAÑO DE AGRUPAMIENTO		
	3	5	7
0.00	377.62 (3.799)	370.57 (3.731)	373.22 (3.726)
0.25	196.22 (1.955)	154.32 (1.502)	131.13 (1.298)
0.50	71.83 (0.691)	50.16 (0.477)	39.79 (0.361)
0.75	30.25 (0.279)	20.97 (0.183)	17.53 (0.142)
1.00	15.24 (0.138)	10.96 (0.083)	9.75 (0.065)
1.25	8.82 (0.074)	7.08 (0.045)	6.83 (0.035)
1.50	5.73 (0.042)	5.17 (0.026)	5.43 (0.021)
1.75	4.19 (0.026)	4.21 (0.017)	4.66 (0.016)
2.00	3.36 (0.017)	3.68 (0.012)	4.16 (0.013)
2.25	2.87 (0.012)	3.31 (0.010)	3.75 (0.012)
2.50	2.56 (0.009)	3.04 (0.009)	3.44 (0.011)
2.75	2.34 (0.007)	2.82 (0.008)	3.16 (0.009)
3.00	2.17 (0.006)	2.62 (0.008)	2.94 (0.009)

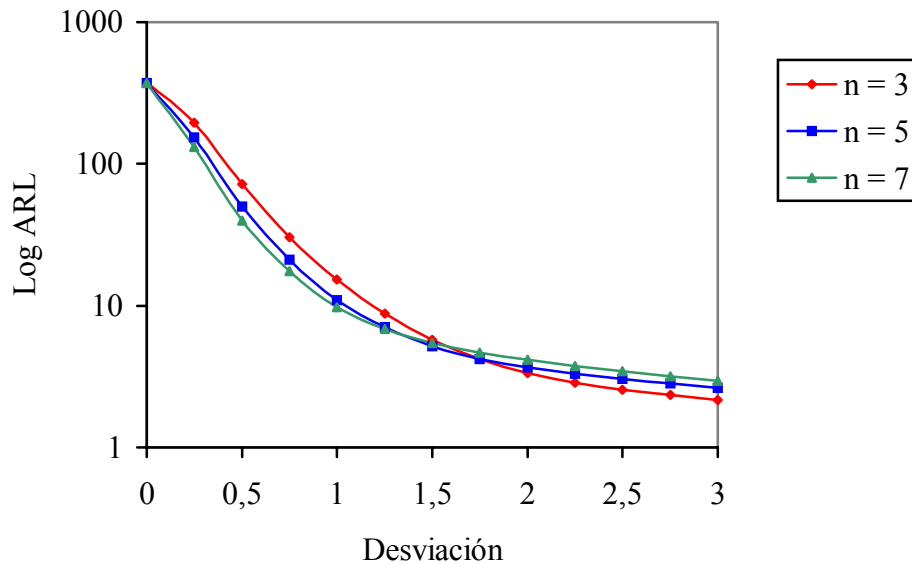
Existe también un rango de desviaciones intermedio para el que la detección de la salida de control se produce más rápidamente si el tamaño de agrupamiento es también intermedio. Por ejemplo, en los resultados obtenidos se observa que cuando  $\delta=1.5\sigma$  el tamaño  $n$  óptimo resulta ser  $n=5$ .

Finalmente, los ARL correspondientes a un gráfico de medias móviles, con tamaño de agrupamiento  $n$  y límites ajustados para obtener un ARL bajo control igual a 370, resultan algo superiores a los correspondientes a un gráfico de medias estándar con tamaño de muestra  $n$ .

Así, cuando la desviación es igual a  $3\sigma$ , en un gráfico de medias estándar con  $n=7$  el ARL es igual a 1, mientras que en el gráfico de medias móviles, en el que también se grafican medias de siete valores, el ARL es igual a 2.94, mayor que el

anterior aunque se trabaja en este último caso con límites de control más estrechos. Esto es consecuencia, tal y como se ha indicado en apartados anteriores, de la estructura de correlaciones existente entre los puntos que se grafican.

FIGURA 5.3- ARL en función del tamaño de agrupamiento  $n=3, 5$  y  $7$ , y de la magnitud de desviación constante de la media del proceso.



### 5.2.3- Rachas en el gráfico de medias móviles

Algunos autores mencionados en la revisión bibliográfica, señalan que la interpretación de los gráficos de medias móviles puede resultar engañosa, porque a medida que aumenta el tamaño de agrupamiento se incrementa el número de pautas o rachas. La apariencia del gráfico en estos casos no se corresponde con ninguna situación que requiera ajustar el proceso. En este sentido, la mayoría de los manuales que describen este procedimiento recomiendan con carácter general no utilizar las señales adicionales de falta de control descritas y analizadas en el capítulo anterior.

En el cálculo de las medias móviles se genera siempre de forma artificial una serie de valores correlacionados, por lo que es de esperar que las rachas aumenten en comparación con lo que ocurre cuando se grafican observaciones independientes. Con el

fin de cuantificar este fenómeno, se ha repetido la simulación del gráfico de medias móviles, variando el tamaño de agrupamiento y para distintas desviaciones en la media.

En cada caso, se incluyen como señales de falta de control las analizadas en el capítulo anterior para el gráfico de medias, y se ha estimado el valor del ARL y la frecuencia relativa de aparición de cada señal. Por tanto, las rachas o pautas consideradas son: dos de tres puntos en la Zona A, siete puntos seguidos a un mismo lado de la línea central y racha de siete puntos seguidos en orden ascendente o descendente. La TABLA 5.4, recoge los valores del ARL para  $n=3, 5$  y  $7$ , con los límites de control en la posición estándar.

TABLA 5.4- ARL del gráfico de medias móviles con rachas, en función del tamaño de agrupamiento y de la desviación en la media.

Entre paréntesis la desviación típica del ARL estimado.

$\delta$ (Ud $\sigma$ )	TAMAÑO DE AGRUPAMIENTO		
	3	5	7
0.00	12.69 (0.133)	9.22 (0.099)	9.42 (0.108)
0.25	9.21 (0.084)	7.06 (0.065)	7.08 (0.066)
0.50	6.83 (0.053)	5.60 (0.043)	5.64 (0.045)
0.75	5.41 (0.037)	4.79 (0.032)	4.82 (0.034)
1.00	4.52 (0.027)	4.19 (0.026)	4.31 (0.028)
1.25	3.81 (0.021)	3.74 (0.021)	3.89 (0.024)
1.50	3.32 (0.017)	3.38 (0.018)	3.61 (0.021)
1.75	2.91 (0.014)	3.13 (0.015)	3.36 (0.019)
2.00	2.66 (0.011)	2.92 (0.014)	3.14 (0.017)
2.25	2.43 (0.009)	2.72 (0.012)	2.97 (0.015)
2.50	2.29 (0.008)	2.57 (0.011)	2.79 (0.014)
2.75	2.15 (0.007)	2.46 (0.010)	2.67 (0.013)
3.00	2.04 (0.007)	2.33 (0.009)	2.54 (0.012)

Comparando esta tabla con la que resulta de tomar como señal de falta de control un punto fuera de los límites (TABLA 5.1), se observa que la consideración de

rachas en el gráfico de medias móviles modifica considerablemente sus propiedades expresadas en función del ARL.

En primer lugar, el valor de este parámetro bajo control disminuye notablemente, siendo mayor este efecto para tamaños de agrupamiento grandes. Esto puede suponer en la práctica un aumento importante en el número de falsas alarmas. Por ejemplo, para  $n=3$  al considerar estas rachas como indicio de salida de control, el intervalo entre dos falsas alarmas coincide en promedio aproximadamente con la toma de 13 muestras, mientras que si se ignoran estas pautas el intervalo está cerca de 377 muestras.

Se observa en este caso que la efectividad es muy similar con  $n=5$  y 7, siendo la de este último tan sólo ligeramente inferior. Por otro lado, cuando la magnitud de desviación es grande, lo que supone un  $\delta > 1.5\sigma$ , el tamaño  $n=3$  presenta una efectividad mayor que la de los otros dos analizados. En general, todos los valores expresados en la TABLA 5.4 son mucho menores que los de la TABLA 5.1, acentuándose la diferencia para desviaciones en la media bajas.

Del apartado anterior se desprendía que el cálculo de medias móviles supone un cambio grande en las propiedades del gráfico, pero este cambio puede ser aún mayor si se tienen en cuenta las rachas que aparecen en el mismo. Este resultado es también consecuencia de la correlación existente entre los puntos registrados. La apariencia normal de un gráfico de medias móviles, estando bajo control o cuando la media ha sufrido una desviación muy pequeña, se caracteriza por la presencia de rachas o pautas, que pueden considerarse en este caso debidas a las causas comunes del procedimiento de control. Estas rachas serán más frecuentes cuanto mayor sea el tamaño de agrupamiento, y en cualquier caso más importantes que en el gráfico de medias, donde se registraban observaciones independientes.

La frecuencia relativa de cada señal en función de la desviación, se recoge en la TABLA 5.5 y su correspondiente FIGURA 5.4, para tamaño de agrupamiento igual a 3.

Se observa en la tabla y figura mencionadas que cuando el proceso está bajo control la frecuencia de aparición de rachas en el gráfico es alta. Así, con tamaño de agrupamiento 3, la racha de siete puntos seguidos a un mismo lado de la línea central y la de siete puntos seguidos en orden ascendente o descendente, constituyen conjuntamente el 84.5% de todas las falsas alarmas que aparecen en el gráfico.

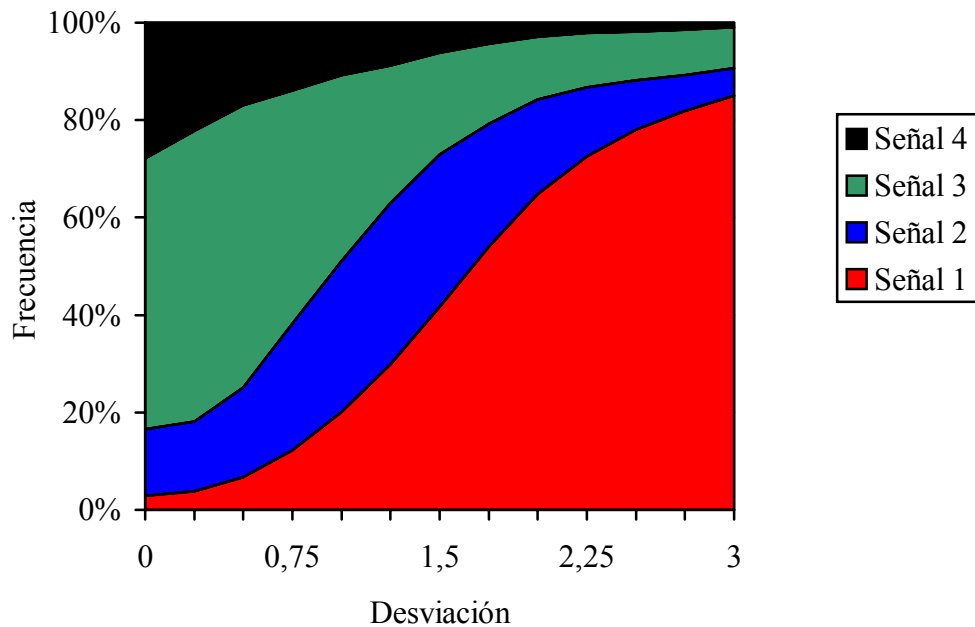
TABLA 5.5- Frecuencia de las rachas en función de la desviación producida en la media. Tamaño de agrupamiento  $n=3$ .

$\delta$ (Ud $\sigma$ )	Señal 1	Señal 2	Señal 3	Señal 4
0.00	2.92	13.60	55.65	27.83
0.25	3.76	14.34	59.74	22.16
0.50	6.74	18.39	57.86	17.01
0.75	12.25	25.99	47.79	13.97
1.00	20.17	30.93	38.11	10.79
1.25	29.87	33.10	28.24	8.79
1.50	41.64	31.23	20.95	6.18
1.75	54.06	25.18	16.41	4.35
2.00	64.83	19.36	12.92	2.89
2.25	72.60	14.11	11.16	2.13
2.50	78.10	10.01	10.10	1.79
2.75	81.83	7.34	9.33	1.50
3.00	84.96	5.74	8.35	0.95

Estos resultados son coherentes con lo mencionado antes para el ARL bajo control, ya que si en esta situación son muy frecuentes las rachas, al ser consideradas como indicio de falta de control en el proceso, suponen una reducción considerable de su valor.

Destaca también la elevada frecuencia relativa de la Señal 4, en comparación con lo que ocurría cuando se incluía en el gráfico de medias con observaciones independientes. Por tanto, la aparición de rachas de este tipo puede estar ligada a una posible falta de independencia entre las observaciones.

FIGURA 5.4- Frecuencia relativa de las rachas en función de la desviación. Tamaño de agrupamiento  $n=3$ .



Tal y como se analiza en el capítulo siguiente, la presencia de correlación puede deberse en otros casos a causas comunes en el proceso, bien sea por el sistema de medida o de ajuste entre dos observaciones sucesivas. En el caso del gráfico de medias móviles, la correlación entre observaciones está implícita en la forma con que se construye, por lo que siempre llevará asociada la aparición de pautas que realmente no reflejan la existencia de causas especiales de variabilidad en el proceso.

Para desviaciones en la media de magnitud grande, se observa que la frecuencia de aparición de un punto fuera de los límites de control aumenta al mismo tiempo que disminuyen las rachas. Así, para una desviación de valor  $3\sigma$  el 84.96% de las señales se deben a esto último.

Con respecto a otros tamaños de agrupamiento  $n$ , los resultados son similares, aunque la importancia de las rachas aumenta conforme lo hace el valor de  $n$ . Se observa que a medida que aumenta el tamaño de agrupamiento lo hace también la frecuencia de rachas y disminuye el número de veces que aparece un punto fuera de los límites de



control. Así, para  $n=2$  la frecuencia de esta señal es igual a 4.23%, mientras que para  $n=7$  se reduce a 0.96%. Aumenta al mismo tiempo la frecuencia de la racha de siete puntos seguidos a un mismo lado de la línea central. Con respecto a la Señal 4, su frecuencia para tamaños pequeños es ligeramente superior a la correspondiente a tamaños grandes.

### **5.3- COMPARACIÓN DE GRÁFICOS DE CONTROL PARA OBSERVACIONES INDIVIDUALES**

De los apartados anteriores se desprende que la utilización de medias móviles implica algunos problemas prácticos, como la aparición de un número excesivo de rachas, así como un cambio de propiedades estadísticas del gráfico, que pueden redundar en una utilización ineficiente. Teniendo esto en cuenta, resulta relevante la consideración de otros métodos para control de procesos con observaciones individuales.

Tal y como se indicó en la introducción, otros gráficos frecuentemente empleados en este contexto son el gráfico de medidas individuales, y los gráficos CUSUM y EWMA. Si el criterio que permite elegir entre estos procedimientos, se basa en la mayor sencillez y facilidad de manejo, el gráfico adecuado resulta ser el de medidas individuales, frente al CUSUM y EWMA que, en general, resultan algo más complicados y difíciles de implantar.

El gráfico de medias móviles sería con este criterio intermedio respecto a los otros, pero presentando los inconvenientes antes mencionados. Sin embargo, aunque la simplicidad es importante, la elección del gráfico a adoptar debe basarse también en sus propiedades estadísticas, con el fin de asegurar un procedimiento efectivo frente a las salidas de control más habituales en el proceso.

Teniendo en cuenta este último punto, se ha abordado el estudio comparativo de los cuatro procedimientos antes mencionados, bajo distintas hipótesis referentes a la

magnitud de desviación constante en la media del proceso. Los parámetros de todos estos gráficos se han ajustado de forma que sus ARL bajo control coincidan aproximadamente con el del gráfico de medidas individuales con límites situados en la posición estándar. Por tanto el ARL de referencia en todos los casos es aproximadamente igual a 370.

El gráfico de medias móviles a incluir en la comparación, se ha diseñado con un tamaño de agrupamiento relativamente bajo, ya que esto suele ser lo habitual en la práctica. Con  $n=3$ , los límites de control se ajustan, por otro lado, en la posición  $L=2.945$ .

Para el gráfico CUSUM, se ha adoptado el valor de referencia estándar  $k=0.5\sigma/\sqrt{n}$ , siendo  $n$  en este caso igual a uno puesto que se va a aplicar sobre medidas individuales. Teniendo en cuenta este  $k$  y que el ARL bajo control deseado es igual a 370, se ha acudido al trabajo de Goel y Wu (1972), donde se obtiene para el intervalo de decisión la magnitud  $h=4.77\sigma$ .

Con respecto al gráfico EWMA, se ha adoptado para el parámetro  $\lambda$  el valor 0.2 recomendado por Hunter (1986). En general, cuanto más pequeño sea este parámetro, mayor será la influencia o peso de observaciones pasadas en la predicción actual, ya que este peso decrece según una progresión geométrica función de  $\lambda$ . Tomando exactamente  $\lambda=0.2$ , la observación correspondiente al instante  $t-10$ , tiene aproximadamente un 11% del peso de la actual.

La posición de los límites de control, que junto con este parámetro proporciona un ARL aproximadamente igual a 370, se ha fijado interpolando en las tablas proporcionadas por Lucas y Saccucci (1990). El EWMA resultante para la comparación, consiste en calcular la predicción correspondiente a cada instante de tiempo, y compararla con los límites situados en:

$$m_0 \pm 2.875 \frac{\sigma}{\sqrt{n}} \sqrt{\frac{\lambda}{2-\lambda}}$$

siendo  $n=1$ , puesto que se aplica a medidas individuales. En el Anejo 1 se desarrolla

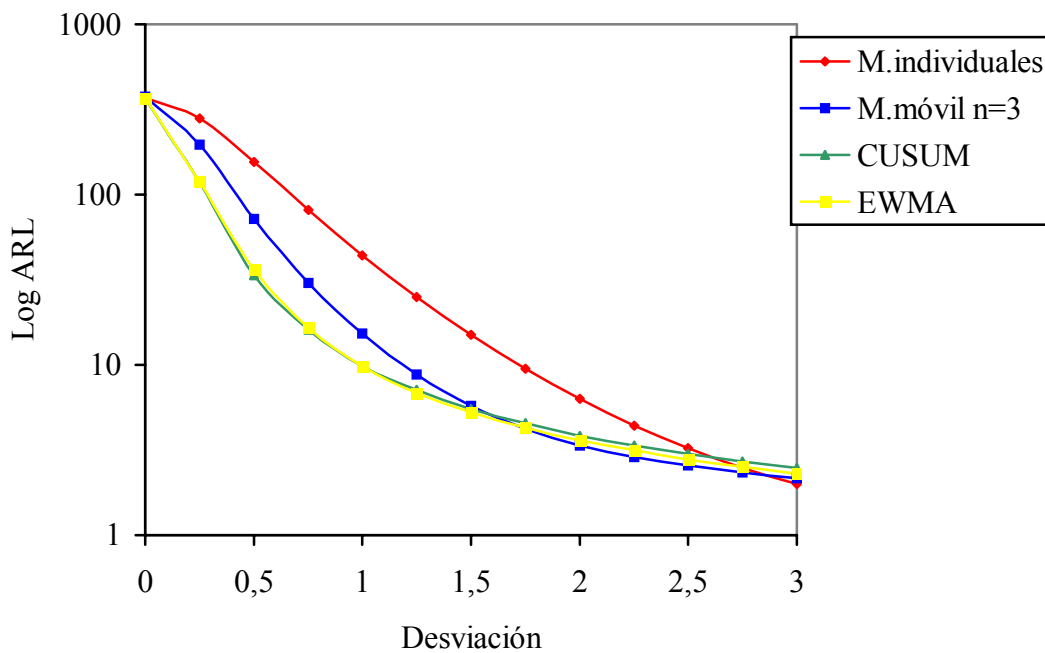
sintéticamente la construcción y diseño de este tipo de gráfico.

La TABLA 5.6 y la correspondiente FIGURA 5.5 recogen los resultados de la simulación de cada uno de estos procedimientos para distintas desviaciones en la media del proceso.

TABLA 5.6- ARL en función de la desviación y del procedimiento de control utilizado para las medidas individuales. Entre paréntesis la desviación típica de la estimación.

GRÁFICOS DE CONTROL				
$\delta$ (Ud $\sigma$ )	Medidas individuales	Media móvil n=3 L=2.945	CUSUM k=0.5 h=4.77	EWMA $\lambda =0.2$ L=2.875
0.00	370.39	377.62 (3.7994)	367.19 (3.6293)	370.25 (3.6681)
0.25	281.14	196.22 (1.9553)	117.4 (1.1069)	119.39 (1.1332)
0.50	155.22	71.83 (0.6910)	33.63 (0.2713)	36.26 (0.3167)
0.75	81.22	30.25 (0.2794)	16.11 (0.1056)	16.56 (0.1177)
1.00	43.89	15.24 (0.1382)	9.82 (0.0522)	9.79 (0.0584)
1.25	24.96	8.82 (0.0737)	7.15 (0.0329)	6.82 (0.0340)
1.50	14.97	5.73 (0.0416)	5.52 (0.0222)	5.25 (0.0233)
1.75	9.46	4.19 (0.0256)	4.54 (0.0163)	4.27 (0.0172)
2.00	6.30	3.36 (0.0168)	3.84 (0.0127)	3.60 (0.0131)
2.25	4.41	2.87 (0.0115)	3.36 (0.0102)	3.16 (0.0106)
2.50	3.24	2.56 (0.0087)	2.99 (0.0085)	2.78 (0.0084)
2.75	2.49	2.34 (0.0073)	2.71 (0.0073)	2.53 (0.0072)
3.00	2.00	2.17 (0.0063)	2.48 (0.0063)	2.30 (0.0061)

FIGURA 5.5- ARL de gráficos para observaciones individuales, en función de la desviación en la media.



Según se observa en la tabla y figura mencionadas, el procedimiento de control óptimo, entendiendo como tal el que detecta más rápidamente una salida de control determinada, depende de la importancia de esta última. Así, para desviaciones constantes de magnitud baja, el gráfico CUSUM resulta ser de todos los procedimientos simulados el más efectivo, lo que se traduce en un valor del ARL correspondiente más pequeño que el de los otros gráficos. Para este mismo rango de desviaciones, el EWMA presenta propiedades similares a las del CUSUM, con un ARL tan sólo ligeramente superior.

Cuando la magnitud de desviación es intermedia, lo que supone en este caso un valor entre  $\sigma$  y  $1,5\sigma$ , el gráfico EWMA resulta ser el procedimiento óptimo. Para cambios en la media superiores a  $1,5\sigma$ , el gráfico de medias móviles presenta los valores del ARL más bajos. Finalmente, cuando esta desviación supera el valor  $3\sigma$ , el gráfico de medidas individuales resulta más efectivo que los otros tres, aunque las diferencias son poco relevantes.

En los resultados expuestos anteriormente se observa que el gráfico de medidas individuales resulta en general el procedimiento menos eficiente, siendo especialmente poco sensible frente a desviaciones de pequeña magnitud. Con respecto a los otros tres gráficos, el CUSUM y EWMA presentan un comportamiento similar, siendo el segundo ligeramente inferior en efectividad al primero, pero en cualquier caso los dos resultan más potentes que el gráfico de medias móviles con  $n=3$ , para desviaciones inferiores a  $1.5\sigma$ .

Para desviaciones superiores a este valor, resulta más efectivo sin embargo agrupar las observaciones en medias móviles, adoptando un tamaño de agrupamiento similar al reflejado en la tabla anterior. Se observa también que el gráfico EWMA resulta ser un procedimiento intermedio entre el CUSUM, y el de medias móviles, estando más próximo al primero.

Los resultados anteriores dependen de los parámetros elegidos para los parámetros de los distintos gráficos. Así, un EWMA con un  $\lambda$  mayor que el estudiado en este trabajo, resulta más próximo al gráfico de media móvil que al CUSUM, y por el contrario con un  $\lambda$  menor, puede incluso resultar más efectivo que el CUSUM para desviaciones bajas.

Por lo tanto, cuando la magnitud de desviación en la media cuya detección precoz resulta relevante es pequeña, el procedimiento de control óptimo puede ser un gráfico CUSUM o EWMA, fijando los parámetros del gráfico correspondiente según lo indicado en el último párrafo. Sin embargo, cuando esta desviación es mayor, un gráfico de medias móviles con tamaño de agrupamiento y límites de control apropiados resulta ser más efectivo.

En la comparación realizada se ha asumido que las observaciones individuales son independientes entre sí. Por lo tanto, los resultados anteriores son válidos siempre que se cumpla esta hipótesis. Sin embargo, el incumplimiento de la condición de independencia se presenta muy a menudo asociado también a los casos en que sólo se

dispone de una medida por intervalo de tiempo. Tal es el caso, por ejemplo, del control de procesos continuos especialmente cuando el intervalo entre observaciones consecutivas es pequeño.

Con el fin de analizar este problema, se ha planteado el análisis abordado en el capítulo siguiente, en el que se cuantifica cómo se modifican las propiedades de los cuatro gráficos antes estudiados con datos correlacionados.

## **CAPÍTULO 6**

### **GRÁFICOS DE CONTROL Y DATOS CORRELACIONADOS**

#### **6.1- INTRODUCCIÓN**

En el capítulo anterior se han estudiado los distintos gráficos aplicables para control on-line en el contexto de medidas individuales. Una de las hipótesis de partida en estos procedimientos es que las observaciones son independientes entre sí. Por tanto, en la discusión de resultados se asume, entre otras cosas, el cumplimiento de esta condición.

Sin embargo, la aplicación de los gráficos de control en situaciones en las que se dispone de una serie de datos individuales va a menudo asociada con la presencia de una pauta sistemática en los mismos. Tal es el caso de la industria química, en la que el carácter continuo de los procesos y los procedimientos de ajuste utilizados pueden provocar que valores consecutivos de una característica de calidad estén correlacionados entre sí.

En la industria de "piezas", como la electrónica, automóvil, etc..., el desarrollo de sistemas de medida sofisticados, que permiten prácticamente cuantificar el valor de una característica crítica en cada unidad que se produce, puede llevar también a esta situación cuando la producción horaria está constituida por un gran número de artículos.

Todos los procedimientos de fabricación presentan cierta inercia frente a la acción de parámetros ambientales o al ajuste de sus condiciones operativas, por lo que la reducción del intervalo de muestreo implica a menudo que una secuencia de observaciones individuales presente autocorrelación.

El empleo de los gráficos de control sin tener en cuenta la pauta sistemática que presentan los datos, puede resultar en una utilización ineficiente de los mismos. Tal y como se vio en el capítulo anterior, el cálculo de medias móviles introduce de forma artificial correlación entre los puntos a graficar, según una pauta que depende a su vez del tamaño de agrupamiento elegido. Esto provoca un cambio de propiedades importante con respecto al gráfico de medias estándar e incrementa de forma notable el número de rachas en el gráfico, lo que puede redundar en una interpretación incorrecta de lo que está pasando realmente en el proceso.

Si la dependencia entre observaciones se presenta siempre de forma sistemática, se debe cuantificar como parte de la variabilidad debida a causas comunes y no como indicio de la presencia de causas especiales o salida de control en el proceso. Por otro lado, en el diseño de todos los gráficos de control se asume independencia entre las observaciones. Esto implica que, tal y como ocurre con las medias móviles, las propiedades del gráfico resultante serán distintas según la pauta y grado de correlación presentes en los datos.

Aunque en la bibliografía sobre el tema se pueden encontrar algunas propuestas para control de procesos cuando se dispone de una serie de medidas correlacionadas, una cuestión previa importante es cuantificar cómo se modifican las propiedades de los gráficos de control en el mismo caso.

Resulta sorprendente la escasez de estudios que aborden este problema, ya que tal y como se indicó en la revisión bibliográfica, los pocos resultados en relación con el mismo se refieren a las propiedades del gráfico CUSUM bajo control o a la estimación de la dispersión del proceso en estas circunstancias.

Teniendo en cuenta lo anterior, se plantean los contenidos de la última parte de la presente tesis. Se pretende evaluar las propiedades del gráfico de medidas individuales, medias móviles, CUSUM y EWMA cuando las observaciones a registrar están correlacionadas según un determinado modelo de partida. Se asume que la característica de calidad sigue un proceso autorregresivo de primer orden, AR(1).



Se considera la efectividad de los gráficos cuando la media del proceso sufre una desviación de magnitud constante, así como su comportamiento estando bajo control, tomando un valor negativo y otro positivo para el coeficiente de correlación entre dos observaciones consecutivas. A continuación, se expone el procedimiento seguido y los resultados para cada gráfico.

## **6.2- CONSIDERACIONES PREVIAS**

También en este caso la magnitud utilizada como medida de la efectividad o del número aproximado de falsas alarmas si el proceso está bajo control, va a ser el ARL. Se trata de cuantificar cómo se modifican las propiedades de los procedimientos más comunes para control on-line de procesos con medidas individuales, a partir de los valores que toma este parámetro en distintas situaciones.

Los gráficos a considerar en el análisis son los estudiados en el capítulo 5, en el que se compararon sus respectivas potencias cuando los datos eran independientes. Se determinará para cada procedimiento el número promedio de observaciones registradas desde que se produce la desviación hasta que aparece la señal. Estando bajo control, se registrará el número promedio de observaciones entre dos falsas alarmas consecutivas.

La determinación analítica del ARL cuando las observaciones a registrar están correlacionadas resulta excesivamente complicada, e incluso en el caso de algún gráfico concreto inabordable. Teniendo esto en cuenta, se ha recurrido para su cálculo a la simulación en ordenador de la serie de valores correlacionados y de cada gráfico, según lo expuesto en el apartado de metodología. El número de repeticiones en cada caso ha sido igual a 10000.

El estudio se ha centrado en un tipo de modelo detectado frecuentemente en el control on-line de características de calidad en las industrias de procesos. Se trata del autorregresivo de primer orden AR(1), en el que el parámetro  $\phi$  oscila normalmente

entre 0.3 y 0.6. Por tanto, la observación correspondiente a un instante  $t$  cualquiera puede expresarse como:

$$X_t = (1-\phi)m_0 + \phi X_{t-1} + a_t$$

siendo  $a_t$  la perturbación aleatoria y  $m_0$  la media del proceso.

Sin pérdida de generalidad, se considerará que las perturbaciones tienen desviación típica  $\sigma_a = 1$  y que la media bajo control es  $m_0 = 0$ . Se ha partido en cada caso de una perturbación simulada  $X_0 = a_0$ , disipándose el efecto de este valor inicial para un número de observaciones suficientemente grande. La serie  $X_t$  generada según lo indicado anteriormente, presenta las características:

$$X_t = \phi X_{t-1} + a_t$$

$$E(X_t) = 0$$

$$\sigma^2 = 1/(1-\phi^2)$$

$$\rho(X_t, X_{t-k}) = \phi^k$$

Para simular una salida de control, se tiene en cuenta que la primera observación después de producirse un cambio de nivel en la media de magnitud  $\delta\sigma$  viene dada por:

$$X_t = \delta\sigma + \phi X_{t-1} + a_t$$

y para  $k \geq 1$ :

$$X_{t+k} = (1-\phi)\delta\sigma + \phi X_{t+k-1} + a_{t+k}$$

que resulta en un nuevo proceso AR(1) de parámetro autorregresivo  $\phi$  y media  $\delta\sigma$ .

Se va a suponer que la varianza  $\sigma^2$  de las observaciones  $X_i$  es conocida y se mantiene constante, por lo que para el cálculo de los límites de control o de los parámetros de cada gráfico se tomará el valor real,  $1/(1-\phi^2)$ .

La hipótesis asumida en el párrafo anterior tiene mucha importancia en el caso de medidas correlacionadas. Normalmente se desconoce el valor de la desviación típica del proceso, teniendo que ser estimada a partir de un número determinado de observaciones iniciales. En el Anejo 1 se recoge una descripción de los procedimientos a seguir en estas situaciones.

Tal y como se indicó en la revisión bibliográfica, la presencia de correlación entre observaciones puede llevar a una estimación de  $\sigma$  errónea, según el método empleado.

Como ejemplo del sesgo en la estimación de  $\sigma$  que puede resultar de la utilización en estas situaciones de las técnicas habituales, se han simulado 100 observaciones de un proceso AR(1) con  $\phi=0.5$ , cuya desviación típica poblacional es  $\sigma=\sqrt{1/(1-\phi^2)}=1.153$ . La desviación típica  $s$  de estos datos ha resultado:

$$s=\sqrt{\frac{\sum_{i=1}^{100}(X_i - \bar{X})^2}{99}}=1.2$$

que es una razonable aproximación al valor real teórico. Por otro lado, la utilización de recorridos móviles, para los que se suele tomar un tamaño de agrupamiento  $n=2$ , conduce en este caso a una estimación igual a:

$$\hat{\sigma} = \frac{\bar{R}}{d_2} = \frac{0.8949}{1.1279} = 0.7934$$

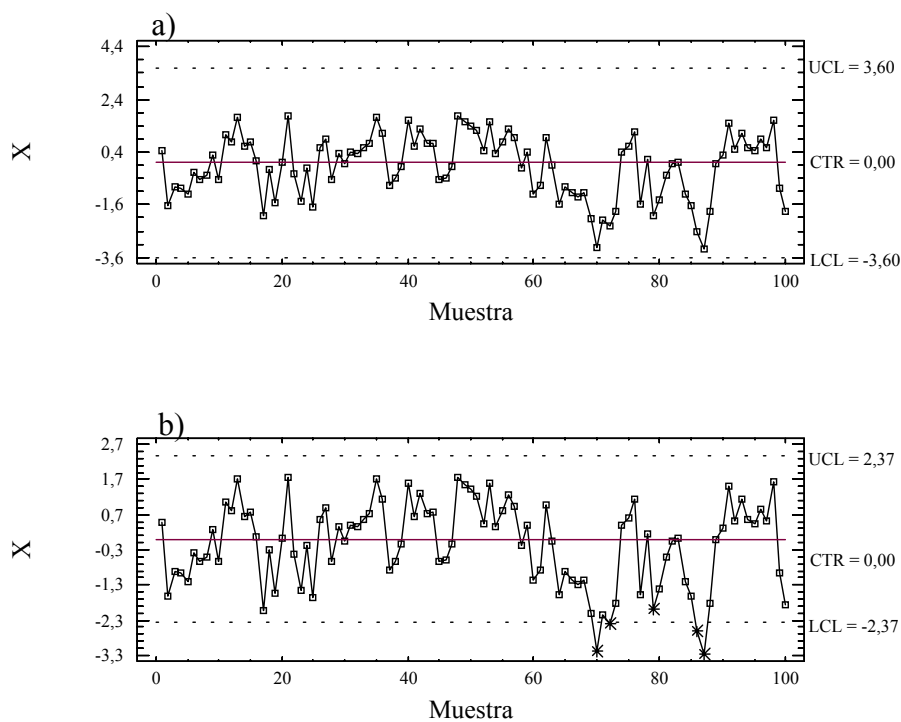
Comparando los dos resultados anteriores, se observa que la estimación obtenida mediante recorridos móviles difiere bastante de la desviación típica real. En este caso concreto, la dispersión del proceso queda infravalorada, lo que en el gráfico de medidas individuales, por ejemplo, redundaría en unos límites de control demasiado estrechos, y por lo tanto en un mayor número de falsas alarmas.

En general, la aplicación de recorridos móviles para estimar la desviación típica en presencia de correlación positiva resultará en un valor inferior al real, puesto que las diferencias entre medidas consecutivas serán más pequeñas que cuando son independientes. Sin embargo, cuando la correlación es negativa aumentan estas diferencias, lo que resulta en una sobreestimación de la dispersión del proceso.

La FIGURA 6.1 recoge el gráfico de medidas individuales aplicado a la serie de los 100 valores simulados según el modelo expuesto anteriormente, y con los límites

estimados con la desviación típica  $s$  y a partir de los recorridos móviles con tamaño de agrupamiento  $n=2$ . Se asume que el proceso está bajo control. Los límites de control en el segundo gráfico son más estrechos que en el primero, y como consecuencia de ello, algunos puntos resultan en falsa alarma. Por otro lado, se detecta en ambos la aparición de rachas por encima y por debajo de la línea central.

FIGURA 6.1- Gráfico de medidas individuales con límites estimados a partir de la a) desviación típica  $s$  y b) a partir de los recorridos móviles.



Además del problema derivado de la estimación incorrecta de la desviación típica del proceso, la falta de independencia entre observaciones puede modificar en otros sentidos sus características operativas. Así, se observa en la FIGURA 6.1 la presencia de rachas, a pesar de que la serie representada es relativamente corta y de que estos proceden de un proceso bajo control. Es de esperar, por tanto, que la correlación modifique considerablemente las propiedades de los gráficos de control y que este cambio dependa a su vez del procedimiento utilizado.

A continuación, se recogen los resultados de la simulación para los gráficos de medidas individuales, medias móviles, CUSUM y EWMA, asumiendo un modelo autorregresivo de primer orden AR(1) para las observaciones y dos valores, uno positivo y otro negativo, para su parámetro  $\phi$ . En todos los casos se supone conocida la desviación típica  $\sigma$  de los datos.

### **6.3- EFECTO DE LA CORRELACIÓN EN EL ARL**

#### **6.3.1- Gráfico de medidas individuales**

En el capítulo anterior se comparó este procedimiento con otros utilizados también en el contexto de observaciones individuales. Los resultados mostraban que es menos efectivo que cualquiera de ellos en la detección de desviaciones constantes en la media inferiores a  $3\sigma$ .

Puesto que la disponibilidad de una sola medida por unidad de tiempo va a menudo asociada a la presencia de correlación, en el presente capítulo se estudian sus propiedades en este contexto. En concreto, se asume que las observaciones siguen un modelo AR(1), tomando para el parámetro  $\phi$  los valores -0.5 y 0.5. Teniendo en cuenta que la correlación aumenta la cantidad de rachas, se considera como única señal de falta de control la aparición de un punto fuera de los límites.

La TABLA 6.1 y la correspondiente FIGURA 6.2 reflejan los resultados de la simulación para distintas desviaciones en la media del proceso, junto con los ARL calculados de forma exacta para el caso de datos independientes.

Se observa que la presencia de correlación no modifica de forma notable las propiedades estadísticas del gráfico de medidas individuales cuando el proceso está bajo control. Así, para  $\phi=0.5$  el ARL no difiere significativamente del correspondiente a datos independientes, y aunque para  $\phi=-0.5$  la diferencia es ligeramente superior, tampoco supone un cambio importante.

Con respecto a las salidas de control simuladas, se observa que la falta de independencia influye algo más en las propiedades del gráfico. Cuando  $\phi > 0$  y la desviación en la media muy pequeña se reduce el tiempo que tarda en producirse la señal de falta de control, mientras que con cambios de nivel superiores a  $0.75\sigma$  ocurre lo contrario.

TABLA 6.1- Efecto de la correlación en el ARL del gráfico de medidas individuales. Entre paréntesis la desviación típica de la estimación.

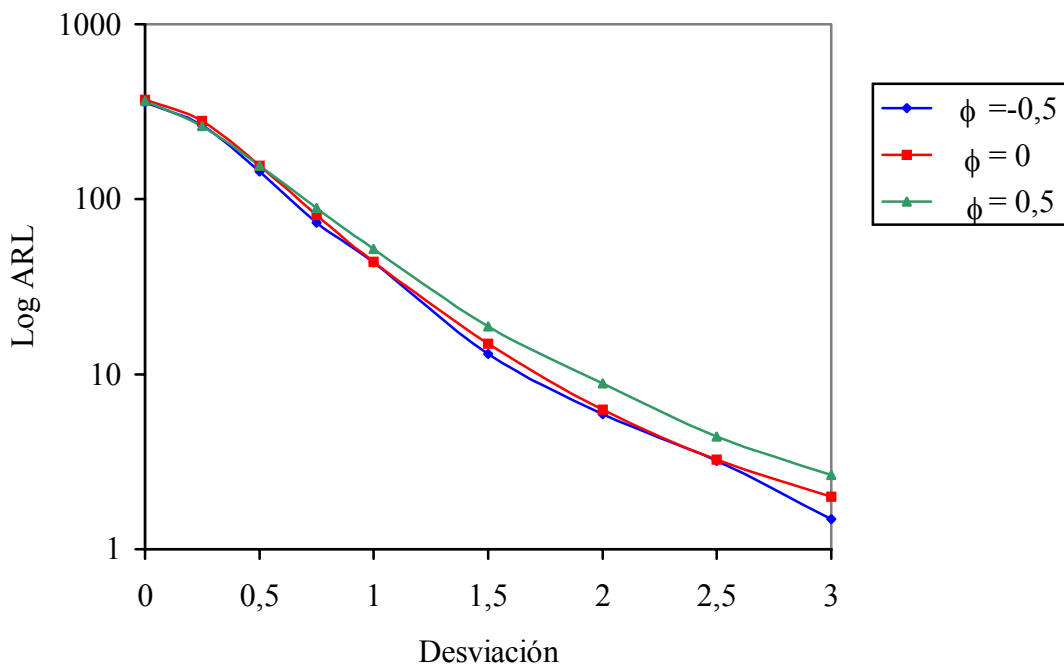
$\delta$ (Ud $\sigma$ )	COEFICIENTE DE CORRELACIÓN		
	-0.5	0	0.5
0.00	358.42 (3.5658)	370.39	364.31 (3.6421)
0.25	263.92 (2.6588)	281.14	262.71 (2.6236)
0.50	144.53 (1.4167)	155.22	155.12 (1.5632)
0.75	73.61 (0.7703)	81.22	88.96 (0.886)
1.00	43.59 (0.4345)	43.89	51.92 (0.5147)
1.50	13.04 (0.1438)	14.97	18.77 (0.1939)
2.00	5.90 (0.0538)	6.30	8.89 (0.0904)
2.50	3.19 (0.0222)	3.24	4.40 (0.0442)
3.00	1.48 (0.0088)	2.00	2.65 (0.0247)

Por otro lado, con un grado de correlación negativa de magnitud intermedia, aumenta la efectividad del gráfico en la detección de las salidas de control simuladas respecto a su comportamiento con observaciones independientes.

La modificación detectada en el valor del ARL se debe a que este parámetro ya no es exactamente igual a  $1/P$  (pto fuera de los límites), como ocurría en el caso de medidas incorrelacionadas.

En el capítulo 5, se vio que asumiendo un proceso markoviano para la variable, la expresión del ARL que se deriva depende del grado de correlación existente entre dos instantes de tiempo consecutivos. Aunque en este caso el modelo es distinto, el valor del ARL difiere también del correspondiente a datos independientes. Cuando  $\phi > 0$ , si una observación se encuentra bajo control aumenta la probabilidad de que con la siguiente ocurra lo mismo, y esto se manifiesta en una disminución de la efectividad. Sin embargo con  $\phi < 0$ , el efecto es el contrario.

FIGURA 6.2- ARL del gráfico de medidas individuales en función de la correlación y la magnitud de desviación en la media.



### 6.3.2- Gráfico de medias móviles

En el gráfico de medias móviles se introduce de forma artificial correlación entre los puntos a registrar. Tal y como se indicó en el capítulo 5, esto provoca un cambio de

propiedades importante respecto al procedimiento de medias estándar.

Adicionalmente, es frecuente su aplicación a observaciones individuales que presentan, como parte de la variabilidad debida a causas comunes, una pauta sistemática que provoca falta de independencia entre las mismas.

A continuación se analiza el efecto que esta circunstancia tiene sobre su comportamiento y efectividad. Se considera un tamaño de agrupamiento  $n$  igual a 3, ya que normalmente se utilizan valores del mismo de este orden de magnitud. Los límites de control están situados en la posición  $m_0 \pm 2.945 \sigma/\sqrt{3}$ , lo que se traduce en un ARL bajo control aproximadamente igual a 370 con datos independientes.

Tal y como se hizo con el gráfico de medidas individuales, se asume que las observaciones se ajustan a un proceso AR(1). Para la correlación negativa se va a considerar un valor más pequeño que el estudiado en el apartado anterior porque a partir de algunas pruebas iniciales se vio que el ARL bajo control y para desviaciones pequeñas aumenta considerablemente en este caso, lo que requiere un tiempo de computación para la simulación excesivamente grande.

La TABLA 6.2 y FIGURA 6.3 recogen los resultados de la simulación en función de la correlación y de la magnitud de desviación producida en la media del proceso.

Se observa en la tabla y figura mencionadas que la presencia de correlación modifica considerablemente las propiedades del gráfico de medias móviles.

En primer lugar, se produce un cambio importante en su comportamiento estando bajo control. Así, con correlación positiva disminuye de forma notable el ARL, lo que se traduce en la práctica en un mayor número de falsas alarmas.

Por otro lado, cuando la correlación es negativa aumenta el valor de este parámetro. En las pruebas realizadas previamente se constató que el incremento que se



produce en el ARL bajo control con un grado de correlación negativa mayor que -0.1 es aún más grande.

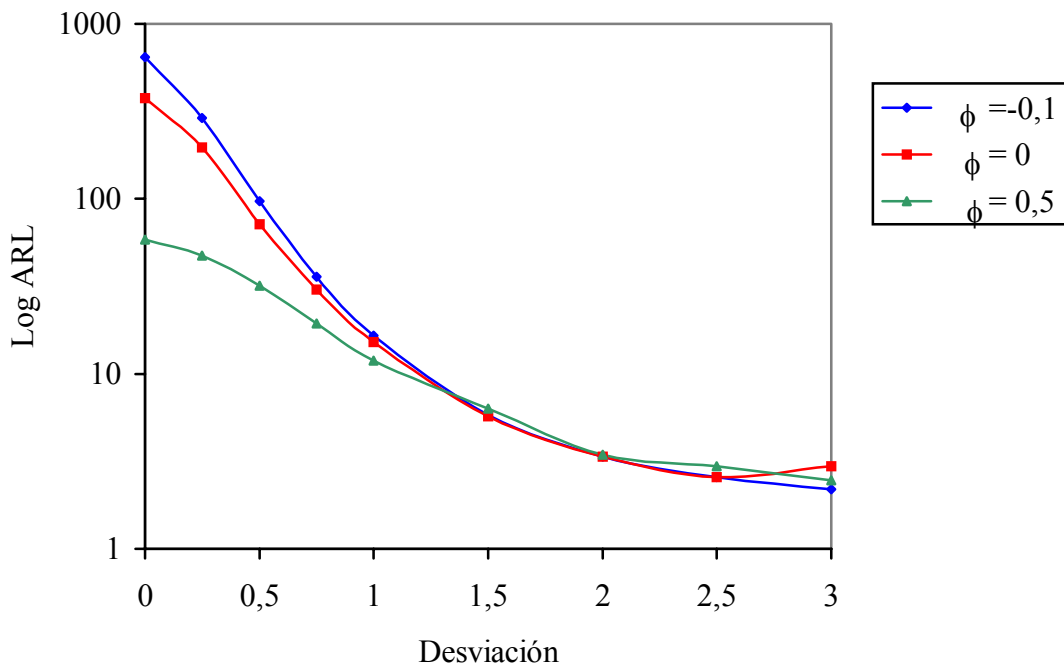
TABLA 6.2- ARL para el gráfico de medias móviles con  $n=3$ , en función de la correlación entre dos observaciones consecutivas y la desviación en la media. Entre paréntesis la desviación típica de la estimación.

$\delta$ (Ud $\sigma$ )	COEFICIENTE DE CORRELACIÓN		
	-0.1	0	0.5
0.00	646.55 (6.6097)	377.62 (3.7994)	58.54 (0.5775)
0.25	290.62 (2.8809)	196.22 (1.9553)	47.42 (0.4908)
0.50	96.71 (0.9746)	71.83 (0.6910)	31.98 (0.3112)
0.75	35.88 (0.3420)	30.25 (0.2794)	19.43 (0.1917)
1.00	16.57 (0.1489)	15.24 (0.1382)	11.91 (0.1208)
1.50	5.83 (0.0410)	5.73 (0.0416)	6.31 (0.0558)
2.00	3.36 (0.0156)	3.36 (0.0168)	3.46 (0.0267)
2.50	2.57 (0.0078)	2.56 (0.0087)	2.96 (0.0155)
3.00	2.19 (0.0061)	2.17 (0.0063)	2.47 (0.0092)

Con respecto a las propiedades del gráfico cuando se produce una desviación en la media de magnitud inferior a  $\sigma$ , la pauta observada en los valores del ARL es semejante a la señalada estando bajo control, resultando inferiores si  $\rho$  es positivo y superiores si  $\rho$  es negativo a los correspondientes al caso de datos incorrelacionados. Para desviaciones superiores a  $\sigma$ , los comportamientos del gráfico en las tres situaciones son cada vez más similares, llegando incluso a invertirse el orden de los valores de los ARL, respecto a la situación constatada para cambios de nivel bajos.

Estos resultados con lógicos puesto que el gráfico de medias móviles aplicado a datos correlacionadas, aunque se construye igual que en el caso de observaciones independientes, resulta un procedimiento con propiedades completamente distintas. Cuando el parámetro  $\phi$  del modelo AR(1) es positivo, se incrementa el número de falsas alarmas, y esto supone, tal y como ocurre con otros procedimientos de control, un aumento de su efectividad frente a desviaciones bajas.

FIGURA 6.3- ARL del gráfico de medias móviles con  $n=3$ , en función de la correlación y de la desviación en la media del proceso.



Al mismo tiempo disminuye la potencia del gráfico frente a desviaciones grandes en la media, aunque en este último caso las diferencias son poco relevantes.

Cuando la correlación es negativa, aunque sea de pequeña magnitud, aumenta notablemente el ARL bajo control, pero paralelamente disminuye la efectividad del gráfico frente a desviaciones bajas en la media. A medida que aumenta la magnitud de

este cambio de nivel, se va igualando su potencia con la correspondiente al mismo procedimiento aplicado a observaciones independientes, siendo de esperar que con un grado de correlación negativa superior al analizado, llegue incluso a superarla.

La justificación de estos resultados puede encontrarse, por un lado, en la modificación que, en presencia de correlación, sufre la probabilidad de que un punto caiga fuera de los límites cuando el anterior no resultó en señal de falta de control. Así, con correlación positiva esta probabilidad es mayor que en el caso de medidas independientes, mientras que con correlación negativa es menor.

Por otro lado, se observa que el cambio detectado en las propiedades del gráfico de medias móviles es mucho más importante que en el gráfico de medidas individuales. Esto es debido a que, además de lo señalado en el párrafo anterior, la falta de independencia modifica también la varianza de las medias móviles graficadas.

En el procedimiento general de construcción de un gráfico de medias móviles, se sitúan los límites de control en la posición  $m_0 \pm L\sigma/\sqrt{n}$ , puesto que  $\sigma/\sqrt{n}$  es la desviación típica de una media móvil calculada a partir de observaciones independientes. A su vez,  $\sigma$  es la desviación típica del proceso y  $n$  el tamaño de agrupamiento adoptado.

Se puede demostrar fácilmente la expresión de la varianza de una media calculada a partir de  $n$  observaciones correlacionadas. Efectivamente si, por ejemplo, el tamaño de agrupamiento es  $n=3$ ,  $\bar{X}_i = (X_i + X_{i-1} + X_{i-2})/3$ , su varianza viene dada por la expresión:

$$\sigma^2_{\bar{X}_i} = \frac{1}{3^2} E[(X_i - m) + (X_{i-1} - m) + (X_{i-2} - m)]^2$$

que desarrollada resulta:

$$\sigma^2_{\bar{X}_i} = \frac{\sigma^2}{3^2} [3 + 4\rho_1 + 2\rho_2]$$

siendo  $\rho_1$  y  $\rho_2$  los coeficientes de autocorrelación de orden 1 y 2 del proceso que siguen los datos.

En el caso que nos ocupa la serie de medidas individuales sigue un modelo AR(1) de parámetro  $\phi$ , siendo su función de autocorrelación simple :

$$\rho(X_t, X_{t-k}) = \phi^k$$

por lo que la varianza de la media de 3 observaciones es igual a:

$$\sigma^2_{\bar{x}_i} = \frac{\sigma}{3} \left[ 1 + \frac{4}{3}\phi + \frac{2}{3}\phi^2 \right]$$

El resultado general de la varianza de una media móvil de n observaciones resulta ser:

$$\sigma^2_{\bar{x}_i} = \frac{\sigma}{n} \left[ 1 + 2 \sum_{k=1}^n \left( 1 - \frac{k}{n} \right) \phi^k \right]$$

Por tanto, si  $\phi > 0$  la varianza de las medias que se grafican es mayor que la que se asume en la construcción del gráfico. Así, los límites de control resultan estrechos en comparación con la variabilidad que presentan las medias registradas. Debido a esto, se incrementa el número de señales de falta de control y disminuye, por tanto, el ARL correspondiente.

Sin embargo, cuando  $\phi < 0$  se produce el efecto contrario, dado que el sumatorio de la expresión anterior resulta negativo con lo que la varianza de las medias es menor que la que se asume al posicionar los límites de control, y por tanto es más improbable la aparición de un punto fuera de ellos.

En el caso que nos ocupa, con  $\phi = 0.5$  la desviación típica de las medias móviles es igual a 0.9027 frente al valor 0.6667 utilizado para la construcción del gráfico, mientras que en el caso de  $\phi = -0.1$  la desviación típica real es 0.5422 y la asumida 0.5802.

Para desviaciones grandes, las consecuencias prácticas del efecto de la correlación en la varianza de las medias no son tan importantes. Se observa que la

efectividad en este caso cambia ligeramente, disminuyendo con  $\phi=0.5$ , y por otro lado aumentando para  $\phi=-0.1$ . Aunque este efecto es justo el contrario del observado con desviaciones pequeñas, es también consecuencia de la modificación que sufre la varianza de las medias según  $\phi$ .

Con  $\phi < 0$  disminuye esta varianza, y al producirse un cambio de nivel importante en la media de las observaciones el porcentaje de la distribución de este estadístico fuera de los límites de control es mayor que con  $\phi > 0$ . Así, para un cambio igual a  $2\sigma$ , con  $\phi=0.5$  esta proporción es aproximadamente igual al 65%, mientras que para  $\phi=-0.1$  es del orden del 71%, por lo que en este último caso en promedio se detectará más rápidamente la desviación.

Una consideración adicional sobre el gráfico de medias móviles aplicado a medidas dependientes es que también se modifica el coeficiente de correlación entre las medias que se grafican.

Si se trata de dos medias móviles consecutivas la expresión del coeficiente con observaciones independientes es:

$$\rho = 1 - \frac{1}{n}$$

Cuando los datos siguen un proceso AR(1) y el tamaño de agrupamiento es  $n$ , el grado de correlación entre dos medias consecutivas resulta ser:

$$\rho = \frac{\sigma^2(\bar{X}_i, \bar{X}_{i-1})}{\sigma_{\bar{X}_i} \sigma_{\bar{X}_{i-1}}}$$

Para el caso de  $n=3$  la covarianza viene dada por la expresión:

$$\sigma^2(\bar{X}_i, \bar{X}_{i-1}) = \frac{\sigma^2}{3^2} [2 + 4\rho_1 + 2\rho_2 + \rho_3]$$

siendo  $\rho_1$ ,  $\rho_2$  y  $\rho_3$  los coeficientes de autocorrelación de orden 1, 2 y 3 del proceso que siguen los datos. Teniendo en cuenta la expresión de estas autocorrelaciones, la covarianza es igual a:

$$\sigma^2 (\bar{X}_i, \bar{X}_{i-1}) = \frac{\sigma^2}{3^2} [2 + 4\phi + 2\phi^2 + \phi^3]$$

En general la covarianza entre dos medias móviles de n observaciones resulta:

$$\sigma^2 (\bar{X}_i, \bar{X}_{i-1}) = \frac{\sigma^2}{n^2} [(n-1) + 2 \sum_{k=1}^n (n-k)\phi^k + \phi^n]$$

por lo que la correlación vendrá dada por la expresión:

$$\rho = 1 - \frac{(1 - \phi^n)}{n + 2 \sum_{k=1}^n (n-k)\phi^k}$$

Volviendo al caso particular de  $n=3$ , este resultado supone que con  $\phi=0.5$  y  $-0.1$ , la correlación entre dos medias consecutivas graficadas es respectivamente igual a 0.84, y 0.617, frente al valor 0.667 correspondiente a observaciones individuales. Para un tamaño de agrupamiento n cualquiera, si  $\phi > 0$  aumenta la correlación entre las medias y con  $\phi < 0$  disminuye.

Este resultado, junto con lo mencionado anteriormente sobre la varianza de los valores registrados, modifica como hemos visto las propiedades del gráfico de medias móviles, respecto a las que tiene cuando los datos son independientes.

### 6.3.3- Gráfico CUSUM

Teniendo en cuenta los resultados obtenidos para el gráfico de medias móviles, es de esperar que también se modifiquen las propiedades y efectividad del gráfico CUSUM según el grado de correlación entre las observaciones.

Con el fin de estudiar este punto, se analiza en este contexto el CUSUM estudiado en el capítulo 5. Por tanto, se considera el valor de referencia  $k=0.5\sigma/\sqrt{n}$  y el intervalo de decisión  $h=4.77\sigma/\sqrt{n}$ , siendo  $n=1$  pues se dispone de un sólo dato por intervalo de tiempo. Esto supone que con medidas independientes el gráfico presenta un ARL bajo control aproximadamente igual a 370.

El procedimiento seguido en este caso es similar al indicado en apartados anteriores. Se ha generado una serie de observaciones según un proceso AR(1), asumiendo un valor positivo y otro negativo para el parámetro  $\phi$ . Se realizó una prueba con  $\phi=-0.5$ , resultando excesivamente grande el tiempo de computación requerido para la estimación del ARL bajo control y con desviaciones pequeñas. Por tanto, se analiza el comportamiento del CUSUM con  $\phi= 0.5$  y  $-0.1$ .

En la TABLA 6.3 y FIGURA 6.4 se recogen los resultados de las simulaciones en función de la correlación entre dos medidas consecutivas y de la magnitud de desviación constante en la media, expresada en unidades de desviación típica  $\sigma$  del proceso.

Se observa en la tabla y figura mencionadas que los cambios detectados en las propiedades estadísticas del gráfico CUSUM, aunque similares a los indicados para el procedimiento de medias móviles, son en este caso de mayor magnitud para alguna situación concreta.

Cuando  $\phi>0$ , lo que supone correlación positiva entre dos medidas consecutivas, se produce una disminución considerable del ARL bajo control. En concreto, este parámetro resulta aproximadamente igual a un 13% del valor correspondiente a observaciones independientes.

Este primer resultado tiene consecuencias prácticas importantes. Supone un número de falsas alarmas mucho mayor, que puede redundar en un sobreajuste del proceso y, por tanto, en una utilización inadecuada del gráfico CUSUM.

Con respecto a sus propiedades con  $\phi<0$ , se observa que en este caso el ARL bajo control sufre un gran aumento. Por tanto, si las observaciones siguen un modelo AR(1) y la correlación entre dos consecutivas es negativa, la cantidad de falsas alarmas disminuye de forma notable.

TABLA 6.3- ARL del gráfico CUSUM en función de la correlación entre observaciones sucesivas y la magnitud de desviación en la media. Entre paréntesis la desviación típica de la estimación.

$\delta$ (Ud $\sigma$ )	COEFICIENTE DE CORRELACIÓN		
	-0.1	0	0.5
0.00	743.67 (7.4732)	367.19 (3.6293)	47.44 (0.426)
0.25	166.06 (1.5618)	117.4 (1.1069)	35.43 (0.3276)
0.50	38.38 (0.3103)	33.63 (0.2713)	23.19 (0.2016)
0.75	16.22 (0.0998)	16.11 (0.1056)	16.13 (0.1206)
1.00	9.84 (0.0483)	9.82 (0.0522)	11.01 (0.0756)
1.50	5.32 (0.0200)	5.52 (0.0222)	5.96 (0.0355)
2.00	3.79 (0.0115)	3.84 (0.0127)	5.08 (0.0219)
2.50	2.94 (0.0080)	2.99 (0.0085)	4.12 (0.0146)
3.00	2.48 (0.0061)	2.48 (0.0063)	2.54 (0.0078)

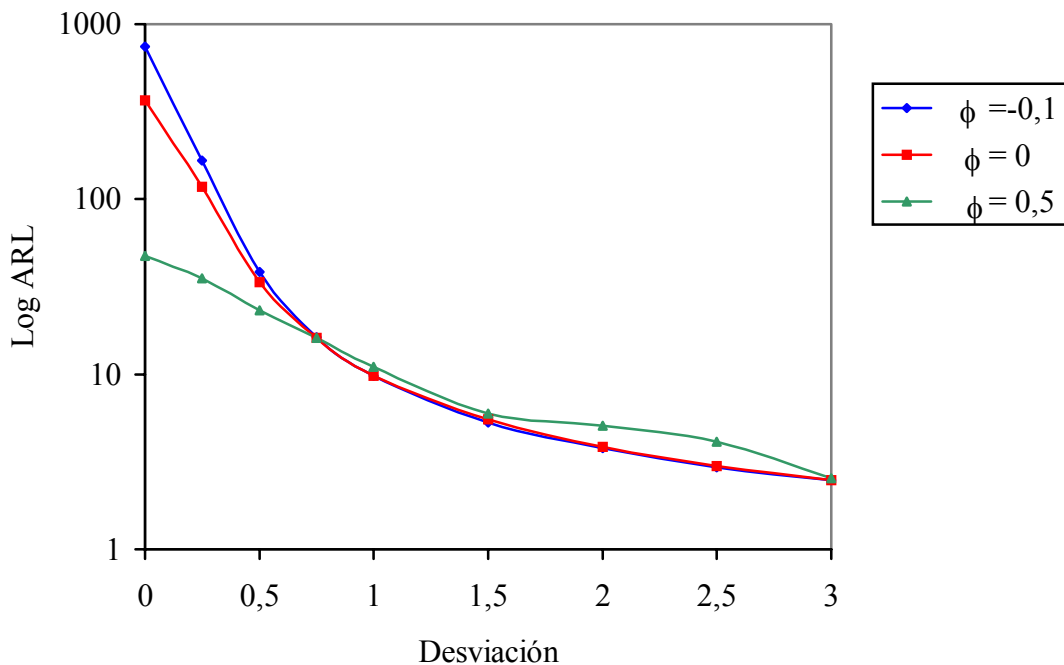
Sin embargo, esto último implica también una menor sensibilidad frente a desviaciones bajas en la media. La diferencia con respecto a sus propiedades con observaciones independientes, disminuye conforme aumenta la magnitud de la desviación. Así, para cambios de nivel superiores a  $1.5\sigma$  la efectividad del gráfico CUSUM no se ve prácticamente modificada por la presencia de correlación negativa de magnitud baja.

Cuando la correlación es positiva el aumento detectado en el número de falsas alarmas va acompañado de una disminución en el tiempo promedio que tarda en detectarse una desviación en la media de magnitud pequeña. A medida que aumenta el



cambio de nivel, el comportamiento del gráfico en presencia de correlación positiva se aproxima al correspondiente a observaciones independientes, llegando incluso a invertirse la situación respecto a la constatada para desviaciones bajas. En efecto, para desviaciones superiores a  $\sigma$ , la efectividad del CUSUM es menor cuando las observaciones siguen un proceso AR(1) con  $\phi > 0$  que cuando los datos están incorrelacionados.

FIGURA 6.4- ARL del gráfico CUSUM en función de la correlación y la magnitud de desviación en la media del proceso.



Por tanto, el comportamiento del gráfico CUSUM bajo control y frente a desviaciones pequeñas cambia considerablemente si las medidas individuales siguen un modelo AR(1), según el signo y valor del parámetro  $\phi$ . La justificación de este resultado se encuentra en la modificación que sufre la varianza del estadístico que se registra en presencia de correlación respecto a la que se asume para la realización del test.

Tal y como se indicó en el apartado del gráfico de medias móviles, la varianza de una media depende de la correlación existente entre las observaciones con las que se calcula. Por tanto, ocurre lo mismo con la varianza de una suma.

Así, la expresión recogida en el apartado anterior, puede aplicarse para determinar la varianza de la suma de I observaciones :

$$S_I = \sum_{i=1}^I X_i$$

$$\sigma^2_{S_I} = \sigma^2 [I + 2 \sum_{i=1}^I (I-i)\phi^i]$$

Teniendo en cuenta que en el gráfico CUSUM se registra la suma acumulada de las diferencias de cada observación respecto del nominal, es de esperar que sus propiedades se vean modificadas. Así, cuando  $\phi > 0$  aumenta la varianza de esta suma acumulada para un instante de tiempo dado, y con  $\phi < 0$  ocurre lo contrario. Como el valor de este estadístico se compara con el intervalo de decisión  $\pm h\sigma/\sqrt{n}$ , siendo  $n=1$  pues se dispone de una serie de observaciones individuales, el resultado será el comentado anteriormente. Para correlaciones positivas aumenta la cantidad de falsas alarmas y la efectividad frente a desviaciones pequeñas, y con correlaciones negativas ocurre justo lo contrario.

Para desviaciones grandes, el efecto del aumento y disminución de la varianza sobre el ARL del gráfico CUSUM es de menor importancia y se manifiesta en otro sentido. La justificación de este resultado es la misma que se indicó en el caso del gráfico de medias móviles. Si disminuye la varianza respecto a la que se asume para el cálculo del intervalo de decisión  $h$  y se produce un cambio de nivel grande, el porcentaje de valores de la distribución del estadístico que se grafica mayores que  $h$  es superior al correspondiente con observaciones independientes, ya que la dispersión de la suma acumulada alrededor de su promedio es menor. Sin embargo, con  $\phi > 0$  el aumento de varianza, implica más dispersión y por lo tanto el efecto contrario.

### 6.3.4- Gráfico EWMA

Este tipo de gráfico es recomendado también para control on-line de procesos en los que se dispone de una sola observación por intervalo de tiempo. Tal y como se indicó en apartados anteriores, esta circunstancia viene asociada frecuentemente con la

falta de independencia en los datos. Situaciones de este tipo se dan, por ejemplo, cuando el intervalo entre medidas consecutivas es muy pequeño, con sistemas de control automático mal calibrados y que operan sobreajustando el proceso, etc...

A diferencia de los otros gráficos estudiados, en el EWMA se registra la predicción obtenida para la característica de calidad, estimada como la media ponderada de la observación y la predicción correspondientes al instante anterior. Por tanto, permite anticiparse en cierto sentido a la medida siguiente, aunque esta predicción sólo es adecuada si el proceso sigue un modelo IMA(1,1). Los límites de control se calculan según la desviación típica de la predicción, asumiendo para la misma su valor asintótico. En el Anejo 1 se recoge una explicación más detallada de este tipo de gráfico.

Las propiedades del gráfico EWMA han sido estudiadas por algunos autores en el caso de observaciones independientes. Sin embargo, no ha sido abordada la misma cuestión en el contexto de datos correlacionados. Teniendo esto en cuenta, se va a determinar su efectividad asumiendo para la serie de observaciones un proceso AR(1), según el procedimiento indicado para los otros gráficos.

En el capítulo 5 se vio que las propiedades de este gráfico son muy similares a las del CUSUM cuando se aplica sobre medidas independientes. Se espera en principio que la modificación del ARL del gráfico EWMA se produzca en el mismo sentido y sea de magnitud similar a la detectada para el CUSUM. Por tanto, se han tomado para  $\phi$  los valores 0.5 y -0.1.

Los parámetros de diseño del EWMA son  $\lambda=0.2$  y  $L=2.875$ . Este procedimiento presenta con observaciones independientes un ARL bajo control aproximadamente igual a 370.

En la TABLA 6.4 y FIGURA 6.5 se recogen los resultados de las simulaciones en función de la correlación entre dos medidas consecutivas y la magnitud de desviación en la media del proceso.

Se observa que la presencia de correlación modifica las propiedades del gráfico EWMA, tal y como ocurre con el gráfico de medias móviles y el CUSUM.

Cuando el proceso está bajo control y  $\phi > 0$  el número de falsas alarmas aumenta considerablemente, mientras que para  $\phi < 0$  disminuye. Las diferencias con respecto a las propiedades del EWMA con datos independientes son ligeramente inferiores a las detectadas en el caso del gráfico CUSUM, y algo superiores a las correspondientes al procedimiento de medias móviles.

Si se produce una salida de control de magnitud pequeña, la pauta observada en las propiedades del gráfico EWMA es también parecida a la de estos dos gráficos. Su potencia disminuye con  $\phi < 0$  y aumenta con  $\phi > 0$ .

Finalmente en el rango de desviaciones grandes el efecto de la correlación se manifiesta en sentido contrario al indicado para cambios en la media pequeños, aunque las diferencias en este caso son poco relevantes.

La justificación de los resultados anteriores, se encuentra también en la manera con que se construye el gráfico EWMA. El valor que se grafica en un instante  $t$  puede expresarse en función de las observaciones obtenidas hasta ese momento, según la expresión:

$$Z_t = \lambda \sum_{j=0}^{t-2} (1-\lambda)^j X_{t-j-1} + (1-\lambda)^{t-1} Z_1$$

Así, para el instante  $t=4$  la varianza de la predicción viene dada por:

$$Z_4 = \lambda X_3 + \lambda (1-\lambda)X_2 + \lambda (1-\lambda)^2 X_1 + (1-\lambda)^3 Z_1$$

$$\sigma_{Z_4}^2 = \lambda^2 \sigma^2 [1 + (1-\lambda)^2 + (1-\lambda)^4 + 2(\rho_1((1-\lambda) + (1-\lambda)^3) + \rho_2(1-\lambda)^2)]$$

siendo  $\rho_1$  y  $\rho_2$  los coeficientes de correlación de orden 1 y 2 del proceso que siguen las observaciones. Como en este caso los datos  $X_i$  siguen un modelo AR(1) de parámetro  $\phi$ , la expresión de la varianza de  $Z_4$  es igual a:

$$\sigma_{Z_4}^2 = \lambda^2 \sigma^2 [1 + (1-\lambda)^2 + (1-\lambda)^4 + 2(\phi((1-\lambda) + (1-\lambda)^3) + \phi(1-\lambda)^2)]$$

TABLA 6.4- ARL del gráfico EWMA en función de la correlación y de la magnitud de desviación.

Entre paréntesis la desviación típica de la estimación.

$\delta$ (Ud $\sigma$ )	COEFICIENTE DE CORRELACIÓN		
	-0.1	0	0.5
0.00	704.31 (7.0018)	370.25 (3.6681)	53.51 (0.5023)
0.25	163.68 (1.5959)	119.39 (1.1332)	36.14 (0.3769)
0.50	40.16 (0.3435)	36.26 (0.3167)	26.32 (0.2258)
0.75	17.11 (0.1197)	16.56 (0.1177)	15.36 (0.1295)
1.00	9.77 (0.0548)	9.79 (0.0584)	11.50 (0.0810)
1.50	5.13 (0.0207)	5.25 (0.0233)	4.89 (0.0344)
2.00	3.53 (0.0118)	3.60 (0.0131)	3.98 (0.0203)
2.50	2.71 (0.0078)	2.78 (0.0084)	2.85 (0.0117)
3.00	2.29 (0.0058)	2.30 (0.0061)	2.30 (0.0071)

En general para un instante t cualquiera, la varianza de  $Z_t$  resulta:

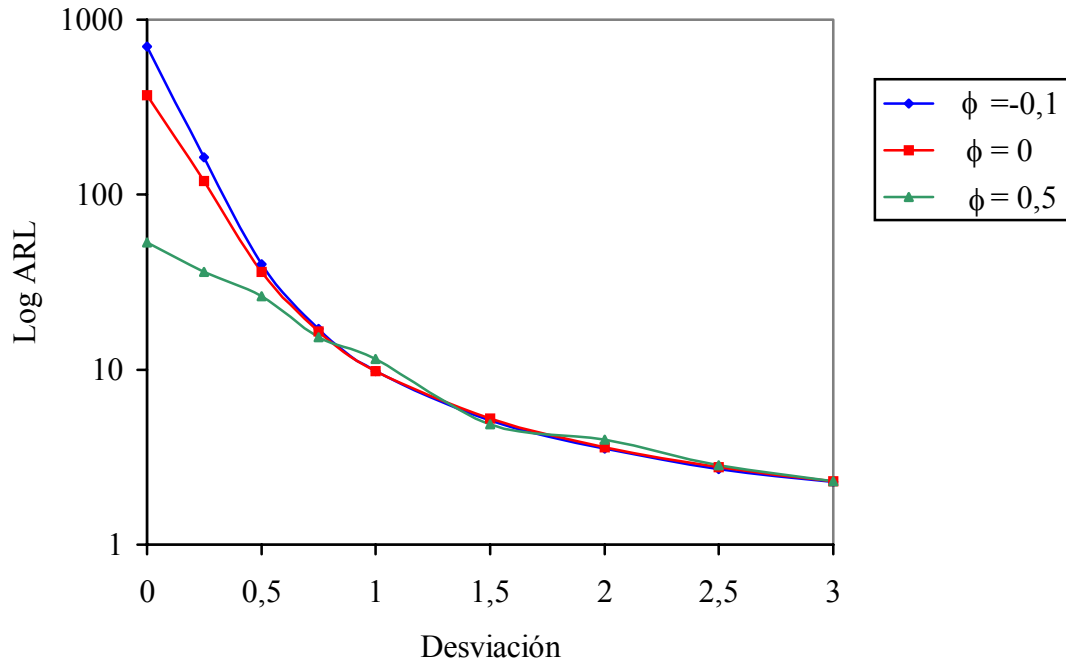
$$\sigma^2_{Z_t} = \lambda^2 \sigma^2 \left[ \sum_{j=0}^{t-2} (1-\lambda)^{2j} + 2 \sum_{j=1}^{t-2} (\phi^j \sum_{k=0}^{t-j-2} (1-\lambda)^{j+2k}) \right]$$

mientras que para la construcción del gráfico se ha asumido el valor:

$$\sigma^2_{Z_t} = \lambda^2 \sigma^2 \left[ \sum_{j=0}^{t-2} (1-\lambda)^{2j} \right] \approx \frac{\lambda}{2-\lambda} \sigma^2$$

Por tanto, cuando  $\phi$  es positivo la varianza de los valores registrados es superior a la calculada para fijar la posición de los límites de control, mientras que con  $\phi$  negativo es inferior. La principal consecuencia de este resultado es la modificación detectada en el valor del ARL bajo control y para desviaciones pequeñas.

FIGURA 6.5- ARL del gráfico EWMA en función de la correlación y de la magnitud de desviación en la media.



Al mismo tiempo, cuando la magnitud de desviación es grande, la probabilidad de que un punto caiga fuera de ellos es mayor si la dispersión de los valores que se grafican es pequeña. Esto implica también una modificación del ARL según el valor del parámetro  $\phi$ , de forma que cuando es negativo y el cambio de nivel importante, aumenta la efectividad respecto a la correspondiente a observaciones independientes, mientras que cuando es positivo ocurre lo contrario.

#### 6.4- CONSIDERACIONES ADICIONALES

Tal y como se indica en apartados anteriores, la presencia de correlación modifica considerablemente las propiedades de los gráficos más utilizados en el contexto de medidas individuales para control on-line de procesos.

Esta modificación depende, a su vez, del procedimiento de control considerado. Así, el gráfico de medidas individuales resulta relativamente robusto frente a la falta de independencia, mientras que la potencia y comportamiento bajo control de los gráficos de medias móviles, CUSUM y EWMA se ven afectados de forma notable en este caso.

El sentido y la magnitud de la variación de la potencia de cada gráfico vienen dados en función del valor de la desviación cuya detección precoz resulta relevante. Para valores pequeños de la misma, el cambio de propiedades es importante en estos tres gráficos, mientras que para desviaciones grandes el comportamiento es muy similar al correspondiente con observaciones independientes.

La primera consecuencia práctica de estos resultados es que la presencia de correlación puede redundar en una utilización ineficiente de los procedimientos habituales para control on-line de procesos.

En relación con esto último se encuentran en la escasa bibliografía sobre el tema algunas propuestas consistentes en modificar los parámetros de diseño del gráfico a adoptar con el fin de compensar en cierto sentido el efecto de la correlación sobre sus propiedades.

Otras propuestas (Berthouex et al., 1974; Alwan y Roberts, 1988) consisten en llevar un gráfico de la serie de observaciones original, que refleja la variabilidad debida a causas comunes, en paralelo con un gráfico de control para los residuos del modelo ARIMA ajustado a los datos, y que servirá para la detección de las causas especiales de variabilidad. Recientemente, Montgomery y Mastrangelo (1991) proponen calcular los residuos a partir de la predicción estimada mediante la aplicación de un alisado exponencial a la serie de medidas individuales, independientemente del modelo que sigan.

La existencia de un modelo que relaciona observaciones correspondientes a distintos instantes de tiempo permite predecir la evolución futura del proceso. Por tanto, no basta con ajustar los parámetros del gráfico a adoptar, de forma que se evite una gran

cantidad de falsas alarmas y se incrementa su sensibilidad frente a ciertos cambios de nivel, o con adoptar otro tipo de procedimiento robusto ante la falta de independencia, sino que se debe utilizar también la estructura de correlaciones existente para asegurar una mejora continua de la calidad del producto.

En estos contextos, la aplicación de procedimientos de control con el único fin de detectar la presencia de causas especiales de variabilidad supone desaprovechar una información valiosa, que podría ser empleada para minimizar la variabilidad respecto del valor óptimo mediante la predicción del valor de la característica de calidad según el proceso que sigan los datos y la realización del ajuste correspondiente, en caso necesario.

Algunos trabajos recientes (Box et al., 1989; Tucker et al., 1991) proponen en este sentido la aplicación de técnicas de control algorítmico (ASPC), que combinan por un lado los enfoques tradicionales del SPC con los del control automático. Teniendo en cuenta la frecuencia con que se presenta la correlación entre las observaciones, es indudable el interés que tiene esta nueva aproximación al problema de control on-line de procesos.



## **CAPÍTULO 7**

### **CONCLUSIONES**

Se sintetizan a continuación los principales resultados obtenidos en el presente trabajo de investigación.

#### **7.1- TAMAÑO Y FRECUENCIA DE MUESTREO EN GRÁFICOS DE MEDIAS**

Cuando se puede asumir un coste de muestreo aproximadamente proporcional al número total de artículos muestreados, el plan de muestreo óptimo para el gráfico de medias estándar depende de la magnitud de desviación constante en la media cuya detección precoz es relevante. Cuando esta desviación es pequeña resulta más efectivo tomar muestras de tamaño relativamente grande espaciadas en el tiempo. Sin embargo, cuando el cambio de nivel es de mayor magnitud, el plan óptimo consiste en tomar muestras de menor tamaño más frecuentemente.

Con respecto a su relación con la capacidad del proceso a controlar, cuando el  $C_p$  es alto el plan óptimo consiste en tomar muestras pequeñas a intervalos cortos de tiempo. El tamaño  $n=5$  recomendado frecuentemente en la práctica, resulta adecuado para la rápida detección de desviaciones que conducen a un 50% de defectuosos cuando el  $C_p$  está comprendida entre 0.6 y 0.8. 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.

#### **7.2-SEÑALES DE FALTA DE CONTROL EN GRÁFICOS DE MEDIAS**

La utilización de señales adicionales de falta de control en el gráfico de medias reduce considerablemente el valor de su ARL bajo control, lo que se traduce en la práctica en un incremento notable del número de falsas alarmas.

Cuando se comparan gráficos con límites ajustados de forma que los ARL bajo control sean similares, se constata que el uso de señales adicionales incrementa la efectividad en el rango de desviaciones inferiores a  $1.67\sigma/\sqrt{n}$ , siendo  $n$  el tamaño de muestra. Si el orden de magnitud de la desviación crítica es superior a este valor, no resulta justificada su utilización pues complican el manejo del gráfico sin mejorar realmente su potencia. En este mismo rango de desviaciones el gráfico CUSUM resulta ser el procedimiento más efectivo.

Cuando se produce una deriva progresiva, las diferencias entre el gráfico de medias estándar, el gráfico con señales adicionales y el CUSUM son poco relevantes. Teniendo esto en cuenta, el criterio a aplicar en este caso para la elección del gráfico de control debe basarse en consideraciones distintas a las estadísticas.

La frecuencia relativa con la que se presentan las diferentes señales de falta de control depende de la magnitud del cambio de nivel en la media. Cuando dicho cambio es pequeño, la esta frecuencia es mayor que la de la señal estándar de falta de control. La frecuencia de la racha de siete puntos seguidos en orden ascendente o descendente es muy baja, por lo que su inclusión en la práctica no está justificada, ya que contribuye al incremento de falsas alarmas pero no a la mejora de la potencia del gráfico de medias.

La utilización de señales adicionales en el gráfico de medias implica la consideración de la información contenida en observaciones pasadas, pero dándoles menor peso según se alejan en el tiempo. Resulta, por tanto, una situación intermedia respecto al gráfico de medias estándar, en el que sólo se tiene en cuenta el último punto graficado, y el gráfico CUSUM, en el que se da el mismo peso a todas las observaciones pasadas. Así, sus propiedades estadísticas son también intermedias respecto a las de estos dos procedimientos.

### **7.3- GRÁFICOS DE CONTROL PARA MEDIDAS INDIVIDUALES**

Las propiedades estadísticas del gráfico de medias móviles dependen del tamaño  $n$  de agrupamiento elegido. Así, con tamaños de agrupamiento grandes, junto con un

menor número de falsas alarmas se produce una pérdida de sensibilidad en el rango de desviaciones pequeñas. Este resultado es consecuencia de la correlación existente entre los valores graficados, puesto que el grado de dependencia entre dos medias móviles  $t$  y  $t-j$  es función de  $n$  y viene dado por la expresión  $1-j/n$ , siendo  $j < n$ .

Para cambios de nivel intermedios la consideración de un  $n$  grande implica una mayor potencia, puesto que la dispersión del estadístico que se registra es menor y por tanto la sensibilidad frente a salidas de control mayor. Sin embargo, frente a desviaciones importantes, el gráfico resulta más efectivo con tamaños de agrupamiento pequeños. Si se elige un  $n$  grande, en las primeras medias que se calculan después de haberse producido la salida de control se promedia un mayor número de observaciones bajo control, con lo que se "amortigua" en cierto sentido el cambio de nivel que ha tenido lugar.

La apariencia de un gráfico de medias móviles estando bajo control o cuando el cambio de nivel es pequeño se caracteriza por la presencia de una gran cantidad de rachas, debidas a la correlación existente entre dichas medias. Estas rachas son tanto más frecuentes cuanto mayor es el tamaño de agrupamiento y, en cualquier caso, mucho más importantes que en el gráfico de medias.

Destaca al respecto la elevada frecuencia relativa de aparición de la racha de siete puntos seguidos en orden ascendente o descendente. Este resultado contrasta con la baja frecuencia registrada para la misma en el gráfico de medias estándar, donde se vio que su consideración como señal de falta de control no estaba justificada desde el punto de vista estadístico.

El procedimiento de control on-line óptimo para medidas individuales depende de la magnitud de desviación en la media cuya detección precoz resulta relevante. Así, cuando ésta es baja resultan más efectivos los gráficos CUSUM y el EWMA, dependiendo este resultado de los parámetros elegidos en su diseño.

El gráfico de medias móviles con tamaño de agrupamiento  $n=3$ , resulta intermedio respecto a los otros procedimientos en este rango de desviaciones, aunque es por otro lado el más efectivo para cambios de nivel superiores. Sin embargo, cuando la desviación es mayor que  $3\sigma$  el gráfico de medidas individuales presenta un ARL mínimo, aunque en este caso las diferencias entre todos los procedimientos comparados son poco relevantes.

#### **7.4- GRÁFICOS DE CONTROL Y DATOS CORRELACIONADOS**

La presencia de correlación entre las observaciones modifica el comportamiento de los gráficos de control para la media. Este cambio depende del signo y magnitud de la dependencia, así como del tipo de gráfico considerado.

En primer lugar, el gráfico de medidas individuales resulta relativamente robusto frente al incumplimiento de la hipótesis de independencia. Sin embargo, en los procedimientos de medias móviles, CUSUM y EWMA el efecto de la correlación se manifiesta modificando notablemente su comportamiento bajo control y frente a desviaciones bajas.

Asumiendo un modelo AR(1) para las observaciones, cuando el parámetro  $\phi$  es mayor que cero aumenta notablemente la cantidad de falsas alarmas al mismo tiempo que la efectividad frente a desviaciones pequeñas en la media. Sin embargo, con  $\phi < 0$  ocurre lo contrario, puesto que disminuyen considerablemente tanto la frecuencia de falsas alarmas como la sensibilidad frente a cambios de nivel bajos.

Para desviaciones grandes en la media, las diferencias respecto al comportamiento con medidas independientes son poco relevantes. Sin embargo, se observa que con correlación negativa aumenta ligeramente la efectividad y con positiva disminuye.

La justificación de estos resultados se encuentra en la modificación que sufre la varianza del estadístico que se grafica cuando se calcula a partir de observaciones

correlacionadas. Cuando  $\phi > 0$  la varianza es mayor que la que se asume para la construcción y diseño del gráfico, mientras que si  $\phi$  es negativo se produce el fenómeno contrario. Esto implica que bajo control y con desviaciones pequeñas aumente o disminuya respectivamente la probabilidad de señales de falta de control. Sin embargo, con desviaciones grandes, la efectividad del gráfico es tanto mayor cuanto menor es la dispersión del estadístico que se registra alrededor de su promedio, lo que justifica la mayor potencia constatada en el caso de autocorrelación negativa.

Teniendo en cuenta lo expuesto, la aplicación de los gráficos de control a datos correlacionados puede redundar en una utilización ineficiente, bien porque aumenta el número de falsas alarmas o porque disminuye la sensibilidad para detectar las salidas reales de control.

Un posible curso de acción a seguir en este caso consiste en modificar los parámetros de diseño del procedimiento correspondiente para amortiguar el efecto de la correlación. La solución a adoptar depende del tipo de gráfico a implantar.

Sin embargo, en estas situaciones parece razonable cambiar el enfoque tradicional del Control Estadístico de Procesos, aprovechando la información sobre el futuro existente en los datos de cada instante con el fin de optimizar la evolución del mismo. Algunos trabajos recientes proponen en este sentido la aplicación de técnicas de control que combinan por un lado los enfoques tradicionales del SPC con los del control automático (APC). Teniendo en cuenta la frecuencia con que se presenta la correlación entre las observaciones, es indudable el interés que tiene esta nueva aproximación al problema de control on-line de procesos.

## CAPITULO 8

### REFERENCIAS BIBLIOGRÁFICAS

Adams, B.M., y Woodall, W.H. (1989), "An Analysis of Taguchi's On-Line Process-Control Procedure Under a Random-Walk Model," Technometrics, Vol. 31, No. 4, pp.401-413.

Adams, B.M., Lowry, C. y Woodall, W.H. (1991), "The Use (and Misuse) of False Alarm Probabilities in Control Chart Design," a aparecer en Frontiers in Statistical Quality Control 4, eds. Lenz, H.J., G.B. Wetherill, y P.Th. Wilrich, P.Th.

AENOR, Norma UNE n. 66-006-74, Nov. 1974.

Alwan, L.C., y Roberts H.V. (1988), "Time-Series Modeling for Statistical Process Control," Journal of Business & Economic Statistics, Vol. 6, No. 1, pp. 87-95.

Amin, R.W., y Ncube M.M. (1991), "Variable Sampling Interval Combined Shewhart-Cumulative Score Quality Control Procedure," Applied Statistics, Vol. 40, No. 1, pp. 1-12.

Aroian, L.A., y Levene, H. (1950), "The Effectiveness of Quality Control Procedures," Journal of the American Statistical Association, 45, pp. 520-529.

Bagshaw, M, y Johnson R.A. (1975), "The Effect of Serial Correlation on the Performance of CUSUM tests II," Technometrics, Vol. 17, No. 1, pp. 73-80.

Barnard, G.A. (1959), "Control Charts and Stochastic Processes," Journal of the Royal Statistical Society, Series B, Vol. 21, No. 2, pp. 239-271.

Berthouex, P.M., Hunter W.H. y Pallesen L. (1978), "Monitoring Sewage Treatment Plants: Some Quality Control Aspects," Journal of Quality Technology, Vol. 10, No. 4, pp. 139-149.

Bissell, A.F. (1969), "Cusum Techniques for Quality Control," Applied Statistics, 18, pp.1-30.

Bissell, A.F. (1978), "An Attempt to Unify the Theory of Quality Control Procedures," Bulletin in Applied Statistics, Vol. 5, pp. 113-128.

Bissell, A.F. (1979), "A Semi-parabolic Mask for CUSUM Charts," The Statistician, Vol. 28, pp. 1-7.

Bissell, A.F. (1984a), "The Performance of Control Charts and Cusums under Linear Trend," Applied Statistics, Vol. 33, No. 2, pp. 145-151. (Corrigendum, Applied Statistics, 35.)

Bissell, A.F. (1984b), "Estimation of Linear Trend from a Cusum Chart or Tabulation," Applied Statistics, Vol. 33, No. 2, pp. 152-157.

Böhm, W., y Hackl P. (1990), "Improved Bounds for the Average Run Length of Control Charts based on Finite Weighted Sums," The Annals of Statistics, Vol. 18, No. 4, pp. 1895-1899.

Bowker, A.H., y Lieberman, G.J. (1972), Estadística para Ingenieros, Englewood Cliffs, N.J.: Prentice/Hall International.

Box, G.E.P., Jenkins G.M. y MacGregor J.F. (1974), "Some Recent Advances in Forecasting and Control, Part II" Applied Statistics, Vol. 23, No. 2, pp. 158-179.

Box, G.E.P. y Jenkins G.M. (1976), Time Series Analysis: Forecasting and Control, Revised Edition, San Francisco:Holden-Day.

Box, G.E.P. (1988), "Quality Improvement: An Expanding Domain for the Application of Scientific Method," Technical Report No. 31, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.

Box, G.E.P., y Kramer, T. (1989), "Statistical Process Control and Automatic Process Control-A Discussion," Technical Report No. 41, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.

Box, G.E.P. (1991), "Understanding Exponential Smoothing: A Simple Way to Forecast Sales and Inventory," Quality Engineering, Vol. 3, No. 4, pp. 561-566.

Box, G.E.P. (1991), "Statistical Process Control: Monitoring and Adjustment," 19th European Meeting of Statisticians, September 2-6, 1991, Barcelona : Spain.

Brook, D., y Evans D.A. (1972), "An Approach to the Probability Distribution of Cusum Run Length," Biometrika, Vol. 59, No. 3, pp. 539-549.

Burr, I.W. (1953), Engineering Statistics and Quality Control, New York:McGraw-Hill.

Burr, I.W. (1967), "The Effect of Non-normality on Constants for  $\bar{X}$  and R Charts," Industrial Quality Control, Vol. 23, No. 11, pp. 563-569.

Butler, J.J. (1983), "Statistical Quality Control with CUSUM Charts," Chemical Engineering, 8, pp. 73-77.

Constable, G.K., y Yen, V. (1991), "Autocorrelation and Control," ASQC Quality Congress Transactions, Milwaukee, WI, May 1991, pp. 375-382.

Craig, R.J. (1984), "Normal Family Distribution Functions: FORTRAN and BASIC Programs," Journal of Quality Technology, Vol. 16, No. 4, pp. 232-236.



Crosier, R.B. (1986), "A New Two-Sided Cumulative Sum Quality Control Scheme," Technometrics, Vol. 28, No. 3, pp. 187-194.

Crowder, S.V. (1987), "A Simple Method for Studying Run Length Distributions of Exponentially Weighted Moving Average Charts," Technometrics, Vol. 29, No. 4, pp. 401-407.

Crowder, S.V. (1989), "Design of Exponentially Weighted Moving Averages and Related Methods," Journal of Quality Technology, Vol. 21, pp. 155-162.

Champ, C.W., y Woodall, W.H. (1987), "Exact Results for Shewhart Control Charts with Supplementary Runs Rules," Technometrics, Vol. 29, No. 4, pp. 393-399.

Champ, C.W., y Woodall, W.H. (1990), "A Program to Evaluate the Run Length Distribution of a Shewhart Control Chart with Supplementary Runs Rules," Journal of Quality Technology, Vol. 22, No. 1, pp. 68-73.

Chiu, W.K. (1973), "Comments on the Economic Design of  $\bar{X}$ -Charts," Journal of the American Statistical Association, Vol. 68, pp. 919-921.

Chiu, W.K., y Wetherhill, G.B. (1974), "A Simplified Scheme for the Economic Design of  $\bar{X}$ -Charts," Journal of Quality Technology, Vol. 6, pp. 63-69.

Davis, R.B., y Woodall, W.H. (1988), "Performance of the Control Chart Trend Rule Under Linear Shift," Journal of Quality Technology, Vol. 20, No. 4, pp. 260-262.

Deming, W.E. (1982), Out of the Crisis: Quality, Production and Competitive Position, Cambridge: Cambridge University Press.

Dudding, B.P., y Jennett W.J.(1944), Quality Control Chart Technique, General Electric, London.

Duncan, A.J. (1956), "The Economic Design of  $\bar{X}$  Charts Used to Maintain Current Control of a Process," Journal of the American Statistical Association, Vol. 51, pp. 228-242.

Duncan, A.J. (1986), Quality Control and Industrial Statistics, Quinta Edición, Homewood, IL: Richard D. Irwin.

Ewan, W.D. (1963), "When and How to Use Cu-Sum Charts", Technometrics, Vol. 5, No 1, pp.1-22.

Ewan, W.D., y Kemp, K.W. (1960), "Sampling Inspection of Continuous Processes with no Autocorrelation Between Successive Results," Biometrika, Vol. 47 (3 y 4), pp. 363-380.

Fearn, T., y Maris, P.I. (1990), "An Application of Box-Jenkins Methodology to the Control of Gluten Addition in a Flour Mill," Technical Report No. 50, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.

Feigenbaum, A.V. (1961), Total Quality Control, New York: McGraw-Hill.

Ford (1984), Statistical Process Control, Código EU880.

Ford (1990), Quality Methods Handbook: Modification of  $\bar{X}$ -R Chart Techniques for Special Situations, Statistical Methods Office, 1/434 Ford Motor Company, Eagleways: Brentwood.

Gitlow, H., Gitlow, S., Oppenheim, A., y Oppenheim, R.(1989), Tools and Methods for the Improvement of Quality, Homewood,IL:Irwin.

Goel, A.L. (1968), A Comparative and Economic Investigation of  $\bar{X}$  and Cumulative Sum Control Charts, Unpublished Ph.D. Thesis, University of Wisconsin, Madison.

Goel, A.L., y Wu, S.M. (1971), "Determination of A.R.L. and a Contour Nomogram for Cusum Charts to Control Normal Mean," Technometrics, Vol. 13, No. 2, pp.221-230

Goldsmith, P.L., y Whitfield, H. (1961), "Average Run Lengths in Cumulative Chart Quality Control Schemes," Technometrics, Vol.13, pp. 11-20.

Grant, E.L., y Leavenworth, R.S. (1980), Statistical Quality Control, New York: McGraw-Hill.

Harrington, H.J. (1990) El Coste de la Mala Calidad, Ediciones Díaz de Santos.

Harris, T.J. y Ross W.R. (1991), "Statistical Process Control Procedures for Correlated Observations," Canadian Journal of Chemical Engineering, Vol. 69, No. 2, pp. 48-57.

Hasen, B.L. (1980), Control de Calidad: Teoría y Práctica, 2ª edición, Editorial Hispano Europea.

Hawkins, D.M. (1981), "A Cusum for a Scale Parameter," Journal of Quality Technology, Vol. 13, No. 4, pp. 228-231.

Hunter, J.S. (1986), "The Exponentially Weighted Moving Average,"Journal of Quality Technology, Vol. 18, No. 4, pp. 203-210.

Hunter, J.S. (1989-90), "A One-Point Plot Equivalent to the Shewhart Chart with Western Electric Rules," Quality Engineering, Vol. 2, No. 1, pp. 13-19.

Ishikawa, K. (1976), Guía de Control de Calidad, New York: UNIPUB.

Jaehn, A.H. (1991), "The Zone Control Chart," Quality Progress, Vol. 24, No. 7, pp. 65-68.

Johnson, N.L. (1961), "A Simple Theoretical Approach to Cumulative Sum Control Charts," Journal of the American Statistical Association, Vol. 56, pp. 835-840.

Johnson, N.L., y Leone F.C. (1977), Statistics and Experimental Design in Engineering and the Physical Sciences, Volume I and II, 2nd ed. New York: Wiley.

Johnson, R.A. y Bagshaw M. (1974), "The Effect of Serial Correlation on the Performance of CUSUM tests," Technometrics, Vol. 16, No. 1, pp. 103-112.

Juran, J.M., Gryna, F.M., Jr., y Bingham, R.S., Jr. (1990), Manual de Control de la Calidad, 2 edición, Ed. Reverté.

Kemp, K.W. (1961), "The Average Run Length of the Cumulative Sum Chart when a V-mask is used," Journal of the Royal Statistical Society B, Vol. 23, pp.149.

Kemp, K.W. (1962), "The Use of Cumulative Sums for Sampling Inspection Schemes," Applied Statistics, Vol. 11,

Kemp, K.W. (1967a), "Formal Expressions which can be used for the Determination of the Operating Characteristic and Average Sample Number of a Simple Sequential Test," Journal of the Royal Statistical Society B, Vol. 29, No. 2, pp. 248-262.

Kemp, K.W. (1967b), "A Simple Procedure for determining Upper and Lower Limits for the Average Sample Run Length of a Cumulative Sum Scheme," Journal of the Royal Statistical Society B, Vol. 29, No. 2, pp. 263-265.

Kemp, K.W. (1971), "Formal Expressions which can be applied to Cusum Charts," Journal of the Royal Statistical Society B, Vol. 33, No. 3, pp. 331-360.

Kleijnen, J.P.C. (1987), Statistical Tools for Simulation Practitioners, New York:Marcel Dekker.

Knappengerger, H.A., y Grandage, A.H.E. (1969), "Minimum Cost Quality Control Tests," AIEE Transactions, 1, pp. 24-32.

Lai, T.L. (1974), "Control Charts Based on Weighted Sums," The Annals of Statistics, Vol. 2, No. 1, pp.134-147.

Lorenzen, T.J., y Vance, L.C., (1986), "The Economic Design of Control Charts: A Unified Approach," Technometrics, Vol. 28, No. 1, pp. 3-10.

Lucas, J.M. (1973), "A Modified 'V' Mask Control Scheme," Technometrics, Vol. 15, No. 4, pp. 833-847.

Lucas, J.M. (1976), "The Design and Use of V-Mask Control Schemes," Journal of Quality Technology, Vol.8, No.1, pp. 1-12.

Lucas, J.M., y Crosier R.B. (1982), "Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM a Head Start," Technometrics, Vol. 24, No. 3, pp. 199-205.

Lucas, J.M. (1982), "Combined Shewhart-CUSUM Quality Control Schemes," Technometrics, Vol. 24, No. 4, pp. 833-847.

Lucas, J.M. (1985), "Counted Data CUSUM's," Technometrics, Vol. 27, No. 2, pp. 129-144.

Lucas, J.M., y Saccucci, M.S. (1990), "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," Technometrics, Vol. 32, No. 1, pp. 1-29.

Lucas, J.M., Saccucci, M.S., y Pahuja, S. (1990), "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," (revised edition) Submitted to the Drexel Working Paper Series.

Lucas, J.M. (1991). 1991 Gordon Research Conference on Statistics in Chemistry and Chemical Engineering, New Hampton, NH, (Comunicación oral).

MacGregor, J.F. (1988), "On-Line Statistical Process Control," Chemical Engineering Progress, Vol. 84, No. 10, pp. 21-31.

McGilchrist, C.A., y Woodyer, K.D. (1975), "Note on a Distribution-free CUSUM Technique," Technometrics, Vol. 17, No. 3, pp. 321-325.

Milton, R.C., y Hotchkiss, R. (1969), "Computer Evaluation of the Normal and Inverse Normal Distribution Functions," Technometrics, Vol.11, No. 4, pp. 817-822.

Montgomery, D.C. (1980), "The Economic Design of Control Charts: A Review and Literature Survey," Journal of Quality Technology, Vol. 12, pp. 75-87.

Montgomery, D.C. (1991), Introduction to Statistical Quality Control, 2nd Edition, New York: John Wiley & Sons.

Montgomery, D.C., y Mastrangelo, C.M. (1991), "Some Statistical Process Control Methods for Autocorrelated Data," Journal of Quality Technology, Vol. 23, No. 3, pp. 179-204.

Munford, A.G. (1980), "A Control Chart based on Cumulative Scores," Applied Statistics, Vol. 29, No. 3, pp. 252-258.

Ncube, M.M., y Woodall, W.H. (1984), "A Combined Shewhart-Cumulative Score Quality Control Chart," Applied Statistics, Vol. 33, No. 3, pp. 259-265.

Nelson, L.S. (1983), "The Deceptiveness of Moving Averages," Journal of Quality Technology, Vol. 15, No. 2, pp. 99-100.

Nelson, L.S. (1988), "Control Charts: Rational Subgroups and Effective Applications," Journal of Quality Technology, Vol. 20, No. 1, pp. 73-75.

Nelson, L.S. (1988), "Calculation of New Limits for  $\bar{X}$ , R Charts when Subgroup Size Is Changed," Journal of Quality Technology, Vol. 20, No. 2, pp. 149-150.

North, W.R.S. (1980), "The Quangle-A Modification of the Cusum Chart," Applied Statistics, Vol. 31, No. 2, pp. 155-158.

Ott, E.R. (1975), Process Quality Control, New York: McGraw-Hill Book Company.

Owen, M. (1989), SPC and Continuous Improvement, Springer-Verlag.

Page, E.S. (1954), "Continuous Inspection Schemes," Biometrika, 41, pp. 100-115.

Page, E.S. (1955), "Control Charts with Warning Lines," Biometrika, Vol. 42, pp. 243-257.

Page, E.S. (1961), "Cumulative Sum Charts," Technometrics, Vol. 3, No. 1, pp. 1-9.

Page, E.S. (1962), "A Modified Control Chart with Warning Lines," Biometrika, Vol. 49, No. 1, pp. 171-176.

Pardo, L., y Valdés, T. (1987), Simulación. Aplicaciones Prácticas en la Empresa, Ed. Díez de Santos, S.A.

Peña, D. (1987), Estadística Modelos y Métodos. 2. Modelos Lineales y Series Temporales. Madrid: Alianza Universidad Textos.

Peña, D., y Prat, A. (1986), Cómo Controlar la Calidad, Manuales IMPI, n. 16, Instituto de la Pequeña y Mediana Empresa Industrial.

Peña, D., Ruiz-Falcó, A. y Díez, D. (1990), "Los Métodos Estadísticos en la Gestión de Calidad de las Empresas Españolas," Trabajos de Estadística, Vol. 5, No. 2, pp. 53-71.

Ramírez, J.G. y Juan, J. (1989), "A Contour Nomogram for Designing CUSUM Charts for Variance," Technical Report No. 33, Center for Quality and Productivity Improvement, University of Wisconsin-Madison.

Ramírez, J.G. (1989), Sequential Methods in Statistical Process Monitoring, Unpublished Ph.D. dissertation, University of Wisconsin-Madison.

Reynolds, M.R. "Approximations to the Average Run Length in Cumulative Sum Control Chart," Technometrics, Vol. 17, No. 1, pp. 65-71.

Reynolds, M.R, Amin, R.W., Arnold, J.C., y Nachlas, J.A. (1988), " $\bar{X}$  Charts With Variable Sampling Intervals," Technometrics, Vol. 30, No. 2, pp. 181-192.

Roberts, S.W. (1958), "Properties of Control Chart Zone Tests," Bell System Technical Journal, Vol. 37, No. 1, pp. 83-114.

Roberts, S.W. (1959), "Control Chart Tests based on Geometric Moving Averages," Technometrics, Vol. 1, No. 3, pp. 239-250.

Robinson, P.B. y Ho, T.Y. (1978), "Average Run Lengths of Geometric Moving Average Charts by Numerical Methods," Technometrics, Vol. 20, No. 1, pp. 85-93.

Romero, R., y Zúnica, L.R. (1987), "Un Modelo Estadístico para el Control del Consumo Energético en una Industria Automovilística," III Jornadas de la Calidad en la Industria Energética, Santander-Mayo 1987, pp. 314-333.



Ryan, T.P. (1989), Statistical Methods for Quality Improvement, New York: John Wiley & Sons.

Saccucci, M.S. y Lucas, J.M. (1990), "Average Run Lengths for Exponentially Weighted Moving Average Control Schemes Using the Markov Chain Approach," Journal of Quality Technology, Vol. 22, No. 2, pp. 154-162.

Saniga, E.M., y Shirland, L.E. (1977), "Quality Control in Practice...A survey," Quality Progress, Vol. 10, No. 5, pp. 30-33.

SAS Institute Inc. (1986), SAS/QC User's Guide ( Version 5 ), Cary, NC: Author.

Shewhart, W.A. (1931), Economic Control of Quality of Manufactured Product, New York: D. Van Nostrand Co., Inc. (Publicado en 1981 por la ASQC, Milwaukee, WI).

Siegmund, D. (1985), Sequential Analysis: Tests and Confidence Intervals, New York: Springer-Verlag.

Statistical Graphics Corporation (1989), STATGRAPHICS User's Guide (Version 4).

Taguchi, G. (1981), On-Line Quality Control during Production, Tokyo, Japan: Japanese Standards Association.

Taguchi, G. (1986), Introduction to Quality Engineering. Designing Quality into Products and Processes, Asian Productivity Organization, New York: UNIPUB.

Tucker, W.T., Faltin, F.W., y Vander Wiel, S. (1991), "Algorithmic Statistical Process Control: An Elaboration," 1991 Gordon Research Conference on Statistics in Chemistry and Chemical Engineering, New Hampton, NH.

- Van Dobben de Bruyn, C.S. (1968), Cumulative Sum Tests-Theory and Practice, New York: Hafner Publishing Company.
- Vance, L.C. (1983), "A Bibliography of Statistical Quality Control Chart Techniques, 1970-1980," Journal of Quality Technology, Vol. 15, pp. 59-62.
- Vardeman, S., y Ray, D. (1985), "Average Run Lengths for CUSUM Schemes When Observations Are Exponentially Distributed," Technometrics, Vol. 27, No. 2, pp. 145-150.
- Vasilopoulos, A.V. y Stamboulis, A.P. (1978), "Modification of Control Chart Limits in the Presence of Data Correlation," Journal of Quality Technology, Vol. 10, No. 1, pp. 20-30.
- Wadsworth, H.M. (1990), Handbook of Statistical Methods for Engineers and Scientists, New York: McGraw-Hill.
- Wadsworth, H.M, Stephens, K.S., y Godfrey, A.B. (1986), Modern Methods for Quality Control and Improvement, New York: John Wiley & Sons, Inc.
- Waldmann, K.H. (1986), "Bounds for the Distribution of the Run Length of One-Sided and Two-Sided CUSUM Quality Control Schemes," Technometrics, Vol. 28, No. 1, pp. 61-67.
- Walker, E., y Philpot, J.W. (1991) "False Signal Rates for the Shewhart Control Chart with Supplementary Run Tests," Journal of Quality Technology, Vol. 23, No. 3, pp. 247-252.
- Weiler, H. (1952), "On the Most Economical Sample Size for Controlling the Mean of a Population," Ann. Math. Statist., Vol. 23, pp. 247-254.
- Western Electric (1956), Statistical Quality Control Handbook, Select Code 700-444.

Wetherill, G.B. y Brown, D.W. (1991), Statistical Process Control: Theory and Practice, Chapman and Hall.

Wheeler, D.J. (1983), "Detecting a Shift in Process Average: Tables of the Power Function for X Charts," Journal of Quality Technology, Vol. 15, No. 4, pp. 155-170.

Woodall, W.H (1983), "The Distribution of the Run Length of One-Sided CUSUM Procedures for Continuous Random Variables," Technometrics, Vol. 25, No. 3, pp. 295-301.

Woodall, W.H. (1984), "On the Markov Chain Approach to the Two-Sided CUSUM Procedure," Technometrics, Vol. 26, No. 1, pp. 41-46.

Woodall, W.H. (1986a), "The Design of CUSUM Quality Control Charts," Journal of Quality Technology, Vol. 18, No. 2, pp. 99-102.

Woodall, W.H. (1986b), "Conflicts Between Deming's Philosophy and the Economic Design of Control Charts," Frontiers in Statistical Quality Control 3, Eds. H.J. Lenz, G.B. Wetherill y P.Th. Wilrich, Würzburg, Germany: Physica Verlag.

Woodall, W.H. (1986c), "Weaknesses of the Economic Design of Control Charts," (letter to the editor), Technometrics, Vol. 28, No. 3, pp. 408-410.

Woodall, W.H., y Adams, B.M. (1991), "A Note on the Statistical Design of CUSUM Charts," 1991 Gordon Research Conference on Statistics in Chemistry and Chemical Engineering- Poster Session, New Hampton, NH.

Zúnica Ramajo, L.R. (1987), "Control Estadístico de Procesos con Media Variable," Homenaje al Profesor Gonzalo Arnaiz Vellando, pp. 725-739, Madrid: Instituto Nacional de Estadística.

## ANEJO 1

### GRÁFICOS DE CONTROL PARA LA MEDIA

#### A.1.1- GRÁFICOS DE MEDIAS DE SHEWHART

Si la característica de calidad toma valores en una escala continua, los gráficos de medias muestrales constituyen la herramienta estándar más utilizada para control de su posición. Básicamente este procedimiento consiste en tomar una muestra de tamaño  $n$  por intervalo de tiempo, y medir en cada artículo el valor de la característica de interés. Si  $X_{ij}$  es la  $i$ -ava observación de la muestra  $j$ , la media muestral se calcula según:

$$\bar{X}_j = \frac{\sum_{i=1}^n X_{ij}}{n}$$

En el procedimiento estándar, una vez registrado este punto en el gráfico, si cae dentro de los límites de control,  $m_0 \pm L\sigma/\sqrt{n}$ , se considera que el proceso está bajo control, y en caso contrario que está fuera de control. En este último caso se procede con la ayuda del gráfico a la identificación de las causas especiales de variabilidad, y a su posterior eliminación.

Desde el punto de vista estadístico, la característica de calidad es una variable aleatoria  $X$ , con una determinada distribución, que en el caso del gráfico de medias muestrales se asume aproximadamente normal, caracterizada por una media  $m_0$  y desviación típica  $\sigma$ . Estos dos parámetros podrán ser conocidos o desconocidos, en cuyo caso se estiman a partir de un número  $k$  inicial de subgrupos, en una fase previa a la implantación del gráfico. Si el proceso está centrado en el nominal, o valor óptimo de calidad,  $m_0$  coincidirá con éste. Cuando se desconoce la posición del proceso, se estima como la media de las medias muestrales según la expresión:

$$\hat{m}_0 = \bar{\bar{X}} = \frac{\sum_{j=1}^k \bar{X}_j}{k}$$

Por otro lado, se utilizan los recorridos  $R_j$  de estos  $k$  subgrupos para estimar la desviación típica del proceso, calculando en primer lugar el recorrido medio:

$$\bar{R} = \frac{\sum_{j=1}^k R_j}{k}$$

y a continuación, obteniendo  $\hat{\sigma} = \bar{R} / d_2$ , siendo  $d_2$  una constante que depende del tamaño de muestra y del número  $k$  de subgrupos tomados en esta fase inicial.

Otra posibilidad es estimar  $\sigma$  a partir de las desviaciones típicas  $s_j$  de cada muestra. En este caso se procede calculando la desviación típica media  $\bar{s}$ , que a su vez se divide por otra constante función también del tamaño de muestra y del número de subgrupos,  $\bar{s} / c_4$ .

A partir de  $\bar{\bar{X}}$  y  $\hat{\sigma}$ , se calculan los límites de control, que en la situación estándar se sitúan a  $\pm 3$  veces la desviación típica de la media muestral ( $\hat{\sigma} / \sqrt{n}$ ) respecto de la línea central  $\bar{\bar{X}}$ . Las dos constantes  $d_2$  y  $c_4$  se encuentran tabuladas junto con el valor  $3 / (d_2 \sqrt{n})$ , que recibe el nombre  $A_2$  en la mayoría de los manuales sobre técnicas de control on-line de procesos. Si las  $k$  muestras caen dentro de los límites, se concluye que el proceso está bajo control por lo que en adelante se utilizan los límites estimados en esta primera fase.

Se presenta a continuación un ejemplo de este tipo de gráfico. El tamaño de muestra es igual a 5, y en cada artículo se mide la desviación de una cierta dimensión respecto del nominal. La TABLA A.1.1 recoge las observaciones correspondientes a 20 muestras. Se desconocen de partida la media y la desviación típica del proceso, por lo que se procederá a su estimación a partir de las muestras.

TABLA A.1.1- Ejemplo de gráfico de medias muestrales.

	3	0	2	-2	8	3	1	4	3	2
	5	3	3	4	7	4	2	2	6	5
	5	9	-4	6	2	8	3	-2	3	6
	6	8	4	3	4	-3	4	-1	-2	3
	5	3	4	5	5	4	-2	3	3	4
$\bar{X}_j$	4.8	4.6	1.8	3.2	5.2	3.2	1.6	1.2	2.6	4
$R_j$	3	9	8	8	6	11	7	6	8	4
	5	3	0	9	0	8	7	3	4	2
	6	7	6	3	5	6	5	5	3	2
	0	3	0	7	4	2	3	6	5	7
	5	5	4	6	2	5	1	0	6	7
	5	0	0	5	2	0	2	7	6	0
$\bar{X}_j$	4.2	3.6	2	6	2.6	4.4	3.6	4.2	4.8	3.6
$R_j$	6	7	6	6	5	8	6	7	6	7

Como el tamaño de muestra es  $n=5$ ,  $d_2$  es igual a 2.326 y  $A_2=0,58$ , se tiene entonces:

$$\text{Línea central} \quad \bar{\bar{X}} = 3.56$$

$$\text{Recorrido medio} \quad \bar{R} = 6.7$$

$$\text{Límites de control} \quad 3.56 \pm 0.58 \times 6.7$$

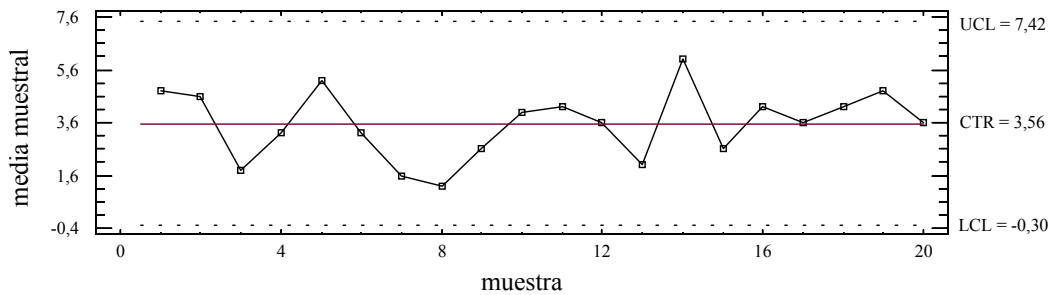
$$\text{LCS: } 7.42$$

$$\text{LCI: } -0.304$$

El gráfico de medias muestrales resultante es equivalente a la realización de un test de hipótesis sobre la posición real de la media del proceso  $m$ . Los límites de control delimitan la región de aceptación de la hipótesis  $m = m_0$ .

Una media muestral  $\bar{X}$  puede caer fuera de los límites de control, incluso cuando el proceso está bajo control; de hecho esto ocurrirá con el 2.73% de las observaciones en el caso estándar  $L=3$ . También puede darse que una muestra procedente de un proceso fuera de control caiga dentro de los límites. Es decir, que hay dos tipos de riesgos inherentes a la operación de un gráfico de control, que se corresponden con los riesgos de primera y segunda especie del contraste clásico de hipótesis.

FIGURA A.1.1- Ejemplo de gráfico de medias muestrales



Además de la señal estándar, un punto fuera de los límites de control, se suele incluir en la práctica habitual otras señales basadas en rachas, con el fin de mejorar la efectividad del gráfico frente a desviaciones de poca magnitud. Su definición implica dividir el gráfico de medias en zonas, según se indica en la FIGURA A.1.2.

Así la Zona A se corresponde con el área comprendida entre los límites de aviso y los de control, situados los primeros normalmente a una distancia  $\pm 2\sigma/\sqrt{n}$  de la línea central. La Zona B es la región situada entre la posición  $m_0 \pm \sigma/\sqrt{n}$  y los límites de aviso. Finalmente la Zona C comprende la zona central del gráfico.

FIGURA A.1.2- Zonas en el gráfico de medias muestrales.

	$m_0 + 3\sigma/\sqrt{n}$
Zona A	
	$m_0 + 2\sigma/\sqrt{n}$
Zona B	
	$m_0 + \sigma/\sqrt{n}$
Zona C	
Zona C	$m_0$
	$m_0 - \sigma/\sqrt{n}$
Zona B	
	$m_0 - 2\sigma/\sqrt{n}$
Zona A	
	$m_0 - 3\sigma/\sqrt{n}$

Algunas de las señales adicionales más frecuentemente utilizadas son (Ishikawa, 1976; Ford, 1984):

- Dos de tres puntos seguidos en la Zona A.
- Una racha de siete puntos seguidos por encima o por debajo de la línea central.
- Demasiados o muy pocos puntos en la Zona C, etc...

### **A.1.2- GRÁFICO DE MEDIDAS INDIVIDUALES**

Es habitual encontrar en la práctica procesos en los que no tiene sentido plantearse la obtención de muestras de un determinado tamaño  $n$ . Este es el caso, por ejemplo, de procesos continuos, como los que se dan en la industria química, o de control de consumos, materia prima, etc...En todas estas situaciones lo más frecuente es disponer de una sola observación por intervalo de tiempo.

Uno de los procedimientos de control a aplicar en estos contextos es el gráfico de medidas individuales, que es un caso particular del gráfico de medias muestrales de Shewhart, para tamaño de muestra  $n=1$ . En este gráfico se registran los valores individuales de la característica de calidad  $X_i$ , y se concluye que el proceso está fuera de control si algún valor cae fuera de los límites de control, o si se presenta alguna de las señales adicionales mencionadas para el gráfico de medias muestrales.

Los valores de la media del proceso  $m_0$  y de la desviación típica  $\sigma$  pueden ser conocidos. En caso contrario se debe proceder a su estimación en una fase previa a la implantación del gráfico, tomando un número  $k$  de observaciones iniciales. A partir de estas, se estima la media como:

$$\bar{X} = \frac{\sum_{i=1}^k X_i}{k}$$



Con respecto a la desviación típica  $\sigma$ , para su estimación se dispone de dos procedimientos distintos. El primero consiste en calcular los recorridos móviles, tomando como tamaño de agrupamiento  $n$ . De esta forma, se dispondrá de  $k-n+1$  recorridos estimados como:

$$R_i = \max(X_{i-n+1}, X_{i-n+2}, \dots, X_i) - \min(X_{i-n+1}, X_{i-n+2}, \dots, X_i)$$

A partir de los  $R_i$ , se estima el recorrido medio  $\bar{R}$ , y tomando el  $d_2$  apropiado,  $\hat{\sigma} = \bar{R}/d_2$ .

Otra posibilidad es utilizar para esta estimación la desviación típica  $s$  de las  $k$  observaciones, que a su vez se calcula mediante la expresión:

$$\hat{\sigma} = s = \sqrt{\frac{\sum_{i=1}^k (X_i - \bar{X})^2}{k-1}}$$

Con cualquiera de las dos estimaciones, los límites de control se calculan como:

$$\text{LCS: } \bar{X} + 3\hat{\sigma}$$

$$\text{LCI: } \bar{X} - 3\hat{\sigma}$$

y se procede a continuación como en el gráfico de medias muestrales anteriormente descrito.

### A.1.3- GRÁFICO DE MEDIAS MÓVILES

Cuando se dispone de una serie de observaciones individuales, una alternativa a la utilización del gráfico de medias individuales, es agrupar las observaciones en medias móviles. La idea básica radica en que como la desviación típica de una media es menor que la de una observación individual, el procedimiento resultante será más sensible, sobre todo frente a desviaciones pequeñas. Por otro lado, en cada punto graficado se tiene también en cuenta parte de la información obtenida en instantes de tiempo anteriores.

En este caso, se adopta un tamaño de agrupamiento  $n$  que determina el número de observaciones consecutivas que se utilizan para calcular cada media. Así, la media  $\bar{X}_i$  se calcula promediando la observación  $X_i$  con las  $n-1$  anteriores:

$$\bar{X}_i = \frac{\sum_{k=i-n+1}^i X_k}{n}$$

Si se dispone inicialmente de  $k$  observaciones, con la expresión anterior se tendrán  $k-n+1$  medias. También se pueden calcular  $k-n+1$  recorridos móviles según lo indicado en el apartado anterior.

A continuación se aplican a estas medias y recorridos las expresiones vistas en el apartado A.1.1 para el gráfico de medias muestrales, para estimar la media y desviación típica del proceso. Una vez calculadas las medias móviles se opera de la forma indicada para este último gráfico.

#### **A.1.4- GRÁFICO CUSUM**

La idea básica de este gráfico radica en que el estadístico que se registra no es el valor observado de la media muestral, sino la suma acumulada de sus desviaciones respecto a la media bajo control. De esta forma en cada punto graficado intervienen todas las observaciones tomadas hasta el momento.

Existen dos versiones distintas del gráfico CUSUM, que resultan en la práctica completamente equivalentes. La primera de ellas utiliza una máscara en  $V$  sobre el gráfico de valores de las sumas acumuladas. Así, para la  $i$ -ava muestra, el estadístico a registrar en el gráfico es:

$$S_i' = \sum_{j=1}^i (\bar{X}_j - m_0)$$

donde se puede estimar  $m_0$  a partir de  $\bar{\bar{X}}$  si se desconoce su valor. Igualmente el valor de  $\sigma$  necesario para construir la máscara, puede estimarse según se ha visto en apartados

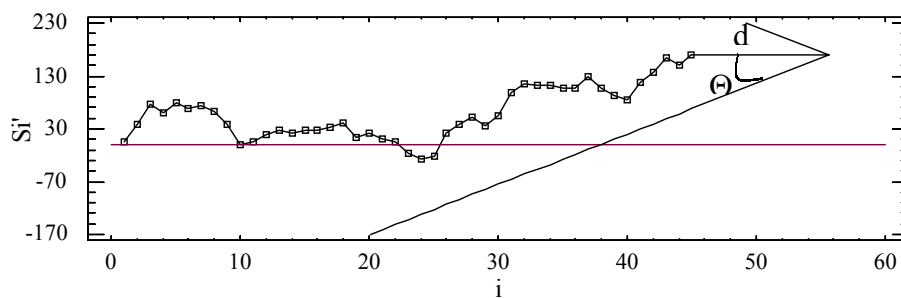
anteriores. La señal de falta de control aparece cuando alguno de los puntos graficados cae fuera de los límites de la máscara en V.

Con el fin de ilustrar este procedimiento, se representa en la FIURA A.1.3 la suma acumulada de las desviaciones individuales procedentes de un proceso químico en batch. En concreto se trata del contenido de un aditivo, TBC, que inicialmente se encuentra centrado en el nominal  $N=150$ , y para el que la desviación típica estimada previamente resulta igual a 19.11. La TABLA A.1.2 recoge los 10 primeros valores observados, junto con la suma acumulada  $S_i'$ .

TABLA A.1.2- Ejemplo de gráfico CUSUM.  $S_0'=0$ .

$i$	$X_i$	$S_i'$
1	155	5
2	183	38
3	188	76
4	135	61
5	168	79
6	139	68
7	155	73
8	141	64
9	125	39
10	111	0

FIGURA A.1.3- Ejemplo de gráfico CUSUM.



Tal y como se observa en la FIGURA A.1.3, la fase de diseño de esta modalidad del gráfico CUSUM consiste básicamente en elegir los valores  $d$  y  $\Theta$  de la máscara en  $V$ , así como el factor de escala apropiado. El valor  $d$  es la distancia entre el último punto registrado y el vértice de la máscara y  $\Theta$  es el ángulo que forman sus “ramas” con la horizontal.

De su valor dependen las propiedades y efectividad del gráfico resultante. Para su elección se debe fijar de antemano el ARL bajo control que se desea y la magnitud de la desviación en la media cuya detección precoz es relevante. Se indica más adelante el procedimiento a seguir. El factor de escala determina el cambio de pendiente en la serie de sumas acumuladas que tiene lugar como consecuencia de una desviación en la media del proceso.

La otra forma de implementar este tipo de gráfico consiste en llevar dos sumas acumuladas definidas como:

$$S_i = \max(0, S_{i-1} + \bar{X}_i - m_0 - k \frac{\sigma}{\sqrt{n}})$$

$$T_i = \max(0, T_{i-1} + \bar{X}_i - m_0 + k \frac{\sigma}{\sqrt{n}})$$

donde con  $S_i$  se detectan desviaciones positivas en la media y con  $T_i$  desviaciones negativas.

Normalmente se toman los valores iniciales  $S_0 = T_0 = 0$ . En este caso la señal de falta de control ocurre cuando  $S_i > h$  ó  $T_i < -h$ . Los parámetros  $k$  y  $h$  se denominan respectivamente valor de referencia e intervalo de decisión, y determinan las propiedades del gráfico resultante.

La TABLA A.1.3 recoge para los mismos datos de antes, los 10 primeros valores de esta suma acumulada, en un gráfico CUSUM con  $k=0.5 \sigma/\sqrt{n}$  y  $h=4.77\sigma/\sqrt{n}$ , valores que suponen un ARL bajo control aproximadamente igual a 370.

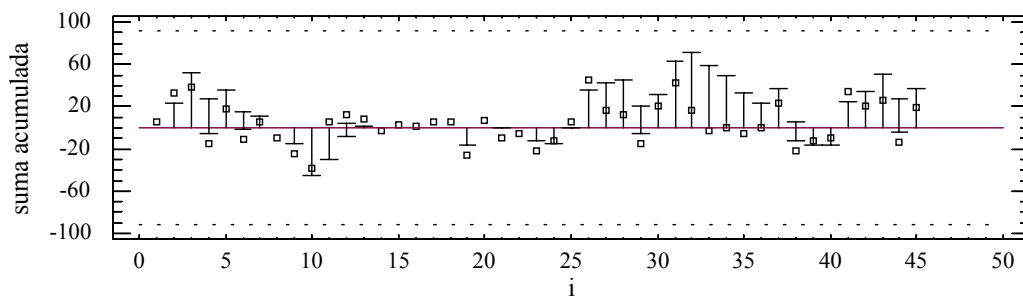
Este procedimiento admite representación gráfica, tal y como muestra la FIGURA A.1.4.

Para fijar los valores de  $k$  y  $h$  se opera de la siguiente manera. En primer lugar se elige el valor deseado para el ARL bajo control y la magnitud  $\delta$  de desviación en la media cuya detección precoz es relevante. Se recomienda tomar un  $k=\delta/2$  lo que resulta en un gráfico CUSUM con un ARL mínimo para una desviación en la media  $\delta\sigma/\sqrt{n}$  y un ARL bajo control dado. El valor  $h$  se determina utilizando alguna de las tablas o ábacos presentes en la bibliografía, de forma que para el  $k$  anterior, el ARL bajo control sea igual que el deseado.

TABLA A.1.3- Ejemplo de gráfico CUSUM.

I	$X_i$	$S_i$	$T_i$
1	155	0	0
2	183	23.45	0
3	188	51.9	0
4	135	27.35	-5.45
5	168	35.8	0
6	139	15.25	-1.45
7	155	10.7	0
8	141	0	0
9	125	0	-15.45
10	111	0	-44.9

FIGURA A.1.4- Ejemplo de gráfico CUSUM.



Si se opta por la versión del CUSUM con máscara en V, los parámetros  $d$  y  $\Theta$  se calculan a partir de  $k$  y  $h$  mediante las expresiones siguientes (Ewan y Kemp, 1960):

$$k = w \tan \Theta$$

$$h = w d \tan \Theta$$

siendo  $w$  el factor de escala del gráfico.

Algunos autores proponen como CUSUM estándar el consistente en tomar  $k=0.5\sigma/\sqrt{n}$  y  $h=5\sigma/\sqrt{n}$ . Si se va a implantar este procedimiento con versión con máscara en V, a partir de estos dos valores se puede calcular  $d$  y  $\Theta$ , adoptando para el factor de escala  $w$  la cantidad  $2\sigma/\sqrt{n}$ , también propuesta como estándar. Así un cambio en la media del proceso de magnitud igual a  $w$  supone un incremento de la pendiente de la serie graficada igual a  $45^\circ$ .

Con las recomendaciones estándar antes señaladas, los parámetros de la máscara en V resultan iguales a  $d=10$  Uds horizontales y  $\Theta=14^\circ$ .

#### **A.1.5- GRÁFICO EWMA**

En el gráfico EWMA el estadístico  $Z_t$  a graficar es la predicción para el instante  $t$ , que se calcula como la media ponderada de la observación  $X_{t-1}$  y la predicción realizada en el instante  $t-1$ . El valor  $X_{t-1}$  puede ser una medida individual o la media muestral correspondiente a ese momento. La ponderación se realiza según el valor del parámetro  $\lambda$ , de forma que cuanto más pequeño sea, tanto mayor resultará la influencia de observaciones alejadas en el tiempo. Un valor recomendado por algunos autores es  $\lambda=2$ .

La expresión a utilizar para el cálculo de la predicción es la siguiente:

$$Z_t = \lambda X_{t-1} + (1-\lambda)Z_{t-1}$$

Para  $Z_1$  se toma normalmente la posición del proceso estando bajo control, que puede estimarse a su vez como la media de  $k$  observaciones iniciales.

En este tipo de gráfico, la predicción correspondiente al instante  $t$  se conoce en  $t-1$ , por lo que permite una cierta anticipación en la toma de decisiones. En algún caso se registra junto a la predicción la serie de valores observados.

Suponiendo que el valor  $Z_t$  es realmente la predicción, se puede expresar también como la diferencia entre la observación  $X_t$  y la perturbación aleatoria correspondiente a ese instante, de la forma:

$$X_t - a_t = \lambda X_{t-1} + (1-\lambda)(X_{t-1} - a_{t-1})$$

y agrupando términos en esta ecuación se llega a:

$$X_t - X_{t-1} = (1-\lambda)a_{t-1} + a_t$$

que es la expresión de un proceso IMA(1, 1) con parámetro  $\theta$  igual a  $1-\lambda$ .

Por tanto, la predicción calculada en un EWMA se corresponde con la de un modelo IMA(1, 1), o proceso de alisado exponencial simple, y que es utilizado frecuentemente para la previsión en el campo de la econometría.

La varianza de  $Z_t$  puede obtenerse desarrollando la expresión que da la predicción, lo que deriva:

$$Z_t = \lambda \sum_{j=0}^{t-2} (1-\lambda)^j X_{t-j-1} + (1-\lambda)^{t-1} Z_1$$

y a partir de ésta:

$$\sigma_{Z_t}^2 = \lambda^2 \sigma^2 \left( \sum_{j=0}^{t-2} (1-\lambda)^{2j} \right)$$

siendo  $\sigma^2$  la varianza de las observaciones. Para el cálculo del sumatorio anterior se puede aplicar la expresión de la suma de un número finito de términos de una progresión geométrica de razón  $(1-\lambda)^{2j}$ :

$$\sigma_{Z_t}^2 = \lambda^2 \sigma^2 \left( \frac{(1-\lambda)^{2(t-2)} - 1}{(1-\lambda)^2 - 1} \right) = \frac{\lambda}{2-\lambda} \sigma^2 [1 - (1-\lambda)^{2(t-2)}]$$

Por tanto, a medida que aumenta  $t$  se aproximará al valor asintótico:

$$\sigma_{Z_t}^2 \approx \frac{\lambda}{2-\lambda} \sigma^2$$

En este tipo de gráfico la señal de falta de control viene dada por la aparición de una predicción fuera de los límites situados en:

$$m_0 \pm L\sigma \sqrt{\frac{\lambda}{2-\lambda}}$$

La desviación típica de las observaciones puede estimarse mediante cualquiera de los procedimientos mencionados en apartados anteriores. El valor de L se escoge, al igual que el de lambda, con el fin de tener un gráfico con unas propiedades y efectividad determinadas.

Como ejemplo de este procedimiento se ha elaborado la FIGURA A.1.5 a partir de las mismas observaciones utilizadas en el apartado anterior para el gráfico CUSUM. La TABLA A.1.4 recoge las 10 primeras observaciones y predicciones, calculadas asumiendo lambda igual a 0.2. La predicción  $Z_1$  es igual al nominal 150.

TABLA A.1.2- Ejemplo de gráfico EWMA.

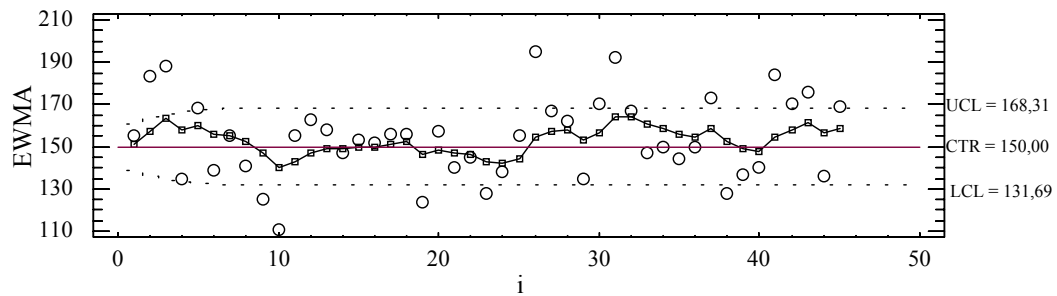
t	$X_t$	$Z_t$
1	155	150
2	183	151
3	188	157.4
4	135	163.52
5	168	164.42
6	139	165.13
7	155	159.91
8	141	158.92
9	125	155.34
10	111	149.27

La posición de los límites se ha ajustado con el fin de tener un ARL bajo control aproximadamente igual a 370, por lo que  $L=2.875$ . Esto resulta en un LCS igual a 168.3 y un LCI situado en 131.7, ya que se asume que el proceso está centrado en el nominal.



En esta figura se incluye la serie de observaciones  $X_t$  original (puntos redondos). Para el diseño de un gráfico EWMA con un parámetro  $\lambda$  y un  $L$  distintos se puede recurrir a las tablas proporcionadas por Lucas y Saccucci (1990), interpolando según el ARL deseado.

FIGURA A.1.5- Ejemplo de gráfico EWMA.



## ANEJO 2 PROGRAMAS UTILIZADOS EN LAS SIMULACIONES

### GENERACIÓN DE PERTURBACIONES ALEATORIAS

∇RESID[□]

[0] RESID  
 [1]  $V \leftarrow W \leftarrow 0 \rho 0 \times I \leftarrow 1$   
 [2]  $E1: V \leftarrow (?2\rho 10*4) \div 10*4$   
 [3]  $\rightarrow (5000 \geq I \leftarrow I + \rho \rho W \leftarrow W, ((-2 \times \oplus V[1]) * 0.5) \times (2, 1) \circ (2 \times \circ V[2])) / E1$   
 [4] ∇

### GRÁFICO DE MEDIAS CON SEÑALES ADICIONALES

#### a) Desviación constante en la media

∇GS4SENALES[□]

[0] ARL1 ← GS4SENALES  
 P;V;E;I;X;Z;N;C;ORAL;M;D;MODA;NP;S;K;ARL;SARL;W  
 [1]  $V \leftarrow 0 \rho 0 \times 0 + W \leftarrow 0 \rho 0 \times K \leftarrow N \leftarrow I \leftarrow 1 + 0 \times \rho S \leftarrow 4 \rho 0 + 0 \times \rho ARL1 \leftarrow 0 \rho 0 + 0 \times \rho E \leftarrow E1, E2, E2, E2, E2$   
 [2] MODA ← 1000 ρ 0  
 [3] SARL ← ARL ← 0  
 [4] M ← P[1]  
 [5] D ← P[2]  
 [6] NP ← P[3]  
 [7] E0:  $V \leftarrow (?2\rho 10*5) \div 10*5$   
 [8]  $\rightarrow (5000 \geq I \leftarrow I + \rho \rho W \leftarrow W, ((-2 \times \oplus V[1]) * 0.5) \times (2, 1) \circ (2 \times \circ V[2])) / E0$   
 [9] F1:  $X \leftarrow (+/(8, M) \rho W[(8 \times M) ? 10000 + I \leftarrow 0]) \div M$   
 [10] E1:  $X \leftarrow 1 \downarrow X, Z \leftarrow D + (+/W[M ? 10000]) \div M + 0 \times I \leftarrow I + 1$   
 [11]  $C \leftarrow 1 \uparrow (((3 \div M * 0.5) < |Z|), (2 \leq +/(2 \div M * 0.5) < -3 \uparrow |X|), (= / \times 1 \downarrow X), (= / \times 2 \downarrow X -- 1 \ominus X)) / t4$   
 [12]  $\rightarrow E[1 + C]$   
 [13] E2:  $\rightarrow ((NP \geq N \leftarrow N + 1) + 0 \times \rho S \leftarrow S + (t4) = C + 0 \times \rho ARL1 \leftarrow ARL1, I) / F1$   
 [14] 'DESVIACION 5.55 TAMANO 5'  $\bar{\phi} D, M$   
 [15] 'FRECUENCIA 55.55 5.55 5.55 5.55'  $\bar{\phi} S \leftarrow S \div NP$   
 [16] 'ARL 55.55'  $\bar{\phi} ARL \leftarrow (+/ARL1) \div NP$   
 [17] E3: SARL ← SARL + (ARL1[K] - ARL) \* 2 + 0 × ρ MODA ← MODA + (ARL1[K] t 1000)  
 [18]  $\rightarrow ((\rho ARL1) \geq K \leftarrow K + 1) / E3$   
 [19] 'MODA 5555555'  $\bar{\phi} (MODA \uparrow / MODA) + 0 \times \rho OARL \leftarrow (OARL \leftarrow ARL1) [\hat{\phi} ARL1]$   
 [20] 'MEDIANA 555555'  $\bar{\phi} (OARL [\lceil NP \div 2 \rceil] + OARL \leftarrow [\lfloor NP \div 2 \rfloor]) \div 2$   
 [21] 'DT 55.55555'  $\bar{\phi} ((SARL \div NP - 1) \div NP) * 0.5$   
 [22] ∇

## b) Desviación progresiva en la media

∇ASC[□]  
[0]NP ASC P;N;I;M;Z;C;X;E;TC  
[1]TC←5ρ0×ρM←P[2]×N←1+0×ρE←E0,E1,E1,E1,E1  
[2]F1:X←(+/(8, M)ρR[(8×M)?10000+I←0])÷M  
[3]E0:X←1↓X,Z←P[1]+(+/R[M?10000])÷M+0×I←I+1  
[4]→E[1+C←1↑(((3÷M\*0.5)<|Z|),(2≤+/(2÷M\*0.5)<-3↑|X|),(=/×1↓X),(=/×2↓X--1⊖X))/ι4]  
[5]E1:→((NP≥N←N+1)+0×ρTC←TC+(I,(ι4)=C))/F1  
[6]TC←TC÷NP  
[7]'DESVIACION 5.55 TAMANO 5'  $\bar{\phi}$  P  
[8]'ARL 55.55 FRECUENCIA 5.55 5.55 5.55 5.55'  $\bar{\phi}$  TC  
[9]∇

## GRÁFICO DE MEDIAS ESTÁNDAR CON LÍMITES DE CONTROL VARIABLES Y DERIVA PROGRESIVA

∇GSC[□]  
[0]NP GSC P;M;N;E;I;X;L;RT;Z;DT  
[1]RT←0×ρM←P[1]×N←1+0×ρL←P[3]+0×ρE←E0,E1  
[2]F1:X←(+/(8, M)ρR[(8×M)?10000+I←0])÷M  
[3]E0:X←1↓X,Z←P[2]+((+/R[M?10000])÷M)+0×I←I+1  
[4]→E[1+((L÷M\*0.5)<|Z|)]  
[5]E1:→((NP≥N←N+1)+0×ρDT←RT\*2+0×ρRT←RT+I)/F1  
[6]'TAMANO 5 DESVIACION 5.55 LIMITES 5.55'  $\bar{\phi}$  P  
[7]'ARL 555.555'  $\bar{\phi}$  RT÷NP  
[8]'DT 5555.5555'  $\bar{\phi}$  (((DT-NP×(RT÷NP)\*2)÷NP-1)\*0.5)÷NP\*0.5  
[9]∇

## GRÁFICO DE MEDIDAS INDIVIDUALES CON OBSERVACIONES CORRELACIONADAS

$\nabla$ MIND[□]  
 [0]NP MIND P;RT;J;N;E;I;AR;ARO;X  
 [1]RT← $2\rho_0+0\times\rho N\leftarrow J\leftarrow 1+0\times\rho E\leftarrow E_0,E_1$   
 [2]AR←W[1?10000]  
 [3]G1:AR←(P[2]×AR)+W[1?10000]  
 [4]J←J+1  
 [5]→(100≥J)/G1  
 [6]ARO←AR  
 [7]F1:I← $0\times\rho X\leftarrow ARO$   
 [8]E0:X←(P[1]×((1-P[2])\*(I>1))÷(1-P[2]\*2)\*0.5)+(P[2]×X)+W[1?10000]+ $0\times\rho I\leftarrow I+1$   
 [9]→E[1+(3÷(1-P[2]\*2)\*0.5)<|X|]  
 [10]E1:→((NP≥N←N+1)+ $0\times\rho RT\leftarrow I*\iota_2$ )/F1  
 [11]'DESVIACION 5.55'  $\bar{\phi}$  P[1]  
 [12]'ARL 55555.555'  $\bar{\phi}$  RT[1]÷NP  
 [13]'DT 55555.5555'  $\bar{\phi}$  (((RT[2]-NP×(RT[1]÷NP)\*2)÷NP-1)÷NP)\*0.5  
 [14]∇

## GRÁFICO DE MEDIAS MÓVILES

### a) Versión estándar

$\nabla$ MEDMOVIL2□  
 [0]RT←MEDMOVIL2 P;N;D;NP;L;RT;E;K;I  
 [1]N←P[1]  
 [2]D←P[2]  
 [3]NP←P[3]  
 [4]L←P[4]  
 [5]RT← $2\rho_0+0\times\rho I\leftarrow 1+0\times\rho E\leftarrow E_0,E_1$   
 [6]F1:K← $0+0\times\rho X\leftarrow W[N?\rho W]$   
 [7]E0:→E[1+(|(+/X←(W[1?ρW]+D),-1↓X)÷N)>L÷N\*0.5+ $0\times\rho K\leftarrow K+1$   
 [8]E1:→(NP≥I←I+1+ $0\times\rho RT\leftarrow RT+K*\iota_2$ )/F1  
 [9]RT←(RT[1]÷NP),(((RT[2]-RT[1]\*2)÷NP)÷NP-1)÷NP)\*0.5  
 [10]'D 5.55 TAMANO 5'  $\bar{\phi}$  D,N  
 [11]'LIMITES 5.555'  $\bar{\phi}$  L  
 [12]'ARL 555.555 DT 5.55555'  $\bar{\phi}$  RT  
 [13]∇

**b) Con señales adicionales de falta de control**

∇MEDMOVIL3□

- [0] MEDMOVIL3 P;N;X;MX;I;K;C;S;Y;E;ARL
- [1]  $N \leftarrow P[1] + 0 \times \rho X \leftarrow 0 \rho 0 + 0 \times \rho MX \leftarrow 0 \rho 0 + 0 \times \rho ARL \leftarrow 2 \rho 0$
- [2]  $K \leftarrow S \leftarrow 0 + 0 \times \rho E \leftarrow E0, E1, E1, E1, E1$
- [3]  $F0: X \leftarrow W[7? \rho W]$
- [4]  $Y \leftarrow 0 \rho I \leftarrow 0$
- [5]  $F1: MX \leftarrow (+ \setminus X \leftarrow W[1? \rho W], -1 \downarrow X) \div i7$
- [6]  $Y \leftarrow Y, MX[N]$
- [7]  $\rightarrow ((\rho Y) < 7) / F1$
- [8]  $E0: MX \leftarrow (+ \setminus X \leftarrow (W[1? \rho W] + P[2]), -1 \downarrow X) \div i7$
- [9]  $Y \leftarrow 1 \downarrow Y, MX[N]$
- [10]  $I \leftarrow I + 1$
- [11]  $C \leftarrow 1 \uparrow (((3 \div N * 0.5) < |MX[N]|), (2 \leq + / (2 \div N * 0.5) < -3 \uparrow |Y|), (= / \times Y), (= / \times Y -- 1 \ominus Y)) / i4$
- [12]  $\rightarrow E[1 + C]$
- [13]  $E1: \rightarrow ((P[3] > K \leftarrow K + 1) + 0 \times \rho S \leftarrow S + (i4) = C + 0 \times \rho ARL \leftarrow ARL + I * i2) / F0$
- [14] 'DESVIACION 5.55'  $\bar{\phi} P[2]$
- [15] 'FRECUENCIA 555.55 555.55 555.55 555.55'  $\bar{\phi} (S \div P[3]) \times 100$
- [16] 'ARL 555.555'  $\bar{\phi} ARL[1] \div P[3]$
- [17] 'DT 55.55555'  $\bar{\phi} (((ARL[2] - (ARL[1] * 2) \div P[3]) \div P[3] - 1) \div P[3]) * 0.5$
- [18] ∇

**c) Versión estándar aplicada a observaciones correlacionadas**

∇MEDMOVIL[□]

- [0] NP MEDMOVIL P;RT;N;K;E;L;AR;ARO;I;Y;J
- [1]  $RT \leftarrow 2 \rho 0 + 0 \times \rho N \leftarrow K \leftarrow 0 + 0 \times \rho E \leftarrow E0, E1$
- [2]  $L \leftarrow (P[2] \div (1 - P[4] * 2) * 0.5) \div P[3] * 0.5$
- [3]  $AR \leftarrow W[1? 10000]$
- [4]  $G1: AR \leftarrow (P[4] \times AR) + W[1? 10000]$
- [5]  $K \leftarrow K + 1$
- [6]  $\rightarrow (100 \geq K) / G1$
- [7]  $ARO \leftarrow AR$
- [8]  $F1: I \leftarrow 0 + 0 \times \rho J \leftarrow 1 + 0 \times \rho Y \leftarrow AR \leftarrow ARO$
- [9]  $F2: Y \leftarrow Y, (AR \leftarrow (P[4] \times AR) + W[1? 10000])$
- [10]  $\rightarrow (P[3] > J \leftarrow J + 1) / F2$
- [11]  $E0: I \leftarrow I + 1$
- [12]  $Y \leftarrow (1 \downarrow Y), (AR \leftarrow (P[1] \times ((1 - P[4]) * (I > 1)) \div (1 - P[4] * 2) * 0.5) + (P[4] \times AR) + W[1? 10000])$
- [13]  $\rightarrow E[1 + L < |(+ / Y \div P[3])|]$
- [14]  $E1: \rightarrow ((NP \geq N \leftarrow N + 1) + 0 \times \rho RT \leftarrow I * i2) / F1$
- [15] 'DESV 5.55 TAMANO 5'  $\bar{\phi} P[1], P[3]$
- [16] 'ARL 55555.555'  $\bar{\phi} RT[1] \div NP$
- [17] 'DT 55555.5555'  $\bar{\phi} (((RT[2] - NP \times (RT[1] \div NP) * 2) \div NP - 1) \div NP) * 0.5$
- [18] ∇

## GRÁFICO CUSUM

### a) Desviación constante en la media

#### a.1) Observaciones independientes

∇GCU[□]

[0]RL←NP GCU P;N;E;I;X;Z;M;RE;PC;H;TC;OA  
[1]PC←2ρ0+0×ρN←I+0×ρE←E0,E1,0ρ(RE←P[3]÷M\*0.5),H←P[4]÷(M←P[2])\*0.5  
[2]RL←0ρ0  
[3]F1:I←0+0×ρZ←2ρ0  
[4]E0:X←P[1]+(+/W[M?10000])÷M+0×I←I+1  
[5]Z←((Z[1]<0),Z[2]>0)×Z←Z+(X+RE),X-RE  
[6]→E[1+(Z[1]<-H)vZ[2]>H]  
[7]E1:PC←PC+(I\*12)  
[8]RL←R  
[9]→(NP≥N←N+1)/F1  
[10]'ARL 555.5555'  $\bar{\phi}$  TC←PC[1]÷NP  
[11]'DT 5.5555'  $\bar{\phi}$  (((PC[2]-NP×TC\*2)÷NP-1)÷NP)\*0.5  
[12]∇

#### a.2) Observaciones correlacionadas

∇GCU[□]

[0]NP GCU P;R;N;J;E;RE;H;I;Z;X;AR;ARO  
[1]R←2ρ0+0×ρN←J←1+0+0×ρRE←P[1]÷(1-P[4]\*2)\*0.5+0×ρE←E0,E1  
[2]H←P[2]÷(1-P[4]\*2)\*0.5+0×ρAR←W[1?10000]  
[3]G1:AR←(P[4]×AR)+W[1?10000]  
[4]J←J+1  
[5]→(100≥J)/G1  
[6]ARO←AR  
[7]F1:I←0+0×ρZ←2ρ0+0×ρAR←ARO  
[8]E0:AR←(P[3]×((1-P[4])\*(I>1)))÷(1-P[4]\*2)\*0.5+(P[4]×AR)+W[1?10000]+0×ρI←I+1  
[9]Z←((Z[1]<0),Z[2]>0)×Z←Z+(AR+RE),AR-RE  
[10] →E[1+(Z[1]<-H)vZ[2]>H]  
[11]E1:R←R+I\*12  
[12]→(NP≥N←N+1)/F1  
[13]'K 5.55 H 5.55 DESV 5.55'  $\bar{\phi}$  3↑P  
[14]'ARL 55555.555'  $\bar{\phi}$  R[1]÷NP  
[15]'DT 55555.5555'  $\bar{\phi}$  (((R[2]-NP×(R[1]÷NP)\*2)÷NP-1)÷NP)\*0.5  
[16]∇

## b) Desviación progresiva y observaciones independientes

∇DCU[□]

[0]NP DCU P;PC;N;E;RE;H;M;I;Z;X

[1]PC←-2ρ0+0×ρN←-I+0×ρE←-E0,E1,0ρ(RE←-P[3]÷M\*0.5),H←-P[4]÷(M←-P[2])\*0.5

[2]F1:I←-0+0×ρZ←-2ρ0

[3]E0:X←-(P[1]×I)+(+/W[M?10000])÷M+0×I←-I+1

[4]Z←-((Z[1]<0),Z[2]>0)×Z←-Z+(X+RE),X-RE

[5]→E[1+(Z[1]<-H)vZ[2]>H]

[6]E1:PC←-PC+(I\*ι2)

[7]→(NP≥N←-N+1)/F1

[8]'ARL 555.5555'  $\bar{\phi}$  PC[1]÷NP

[9]'DT 5.5555'  $\bar{\phi}$  (((PC[2]-NP×TC\*2)÷NP-1)÷NP)\*0.5

[10]∇

## GRÁFICO EWMA

### a) Aplicado a observaciones individuales e independientes

∇EWMA[□]

[0]NP EWMA P;RT;N;E;L;I;Z;X

[1]RT←-2ρ0+0×ρN←-I+0×ρE←-E0,E1

[2]L←-P[3]×(P[3]÷(2-P[3]))\*0.5

[3]F1:Z←-I←-0

[4]E0: I←-I+1

[5]X←-W[1?10000])

[6]Z←-(P[2]×X)+(1+P[2])×Z

[7]→E[1+L<|Z]

[8]E1:→ ((NP≥N←-N+1)+0×ρRT←-RT+I\*ι2)/F1

[9]'DESV 5.55 LAMBDA 5.55 LIMITES 5.55'  $\bar{\phi}$  P

[10]'ARL 555.5555'  $\bar{\phi}$  RT[1]÷NP

[11]'DT 5.5555'  $\bar{\phi}$  (((RT[2]-NP×(RT[1]÷NP)\*2)÷NP-1)÷NP)\*0.5

[12]∇

**b) Aplicado a observaciones individuales correlacionadas**

$\nabla$ EWMA[□]  
 [0]NP EWMA P;RT;N;E;L1;Z;Y;I;J;AR;ARO  
 [1]RT←-2ρ0+0×ρJ←-0+0×ρN←-1+0×ρE←-E0,E1  
 [2]L1←-(P[4]÷((1-P[2]\*2)\*0.5))×(P[3]÷(2-P[3]))\*0.5  
 [3]AR←W[1?10000]  
 [4]G1:AR←-(P[2]×AR)+W[1?10000]  
 [5]J←J+1  
 [6]→(100≥J)/G1  
 [7]ARO←AR  
 [8]F1:I←Z←-0+0×ρAR←ARO  
 [9]E0:AR←-(P[1]×((1-P[2])\*(I>1))÷(1-P[2]\*2)\*0.5)+(P[2]×AR)+W[1?10000]+0×ρI←-I+1  
 [10]Z←-(P[3]×AR)+(1-P[3])×Z  
 [11] →E[1+L1<|Z]  
 [12]E1:→((NP≥N←N+1)+0×ρRT←-RT+I\*12)/F1  
 [13]'DESV 5.55 LAMBDA 5.55 LIMITES 5.55'  $\bar{\phi}$  P[1],P[2],P[3]  
 [14]'ARL 55555.555'  $\bar{\phi}$  RT[1]÷NP  
 [15]'DT 55555.5555'  $\bar{\phi}$  (((RT[2]-NP×(RT[1]÷NP)\*2)÷NP-1)÷NP)\*0.5  
 [16]∇