



Universidade de Vigo

DESARROLLO DE UNA APLICACIÓN PARA GRÁFICOS DE CONTROL
DE PROCESOS INDUSTRIALES

POR
MIGUEL A. FLORES SÁNCHEZ

Proyecto presentado para el cumplimiento de los requisitos para culminar el Máster
en Técnicas Estadísticas de la Universidad de La Coruña

Directores:
Dr. Salvador Naya
Dr. Rubén Fernández Casal

Julio 2013

Resumen

El presente trabajo tiene por objetivo crear una aplicación para el control estadístico de la calidad basado en un entorno gráfico cuyo motor de cálculo sea el lenguaje R. Se implementan los gráficos de control de calidad más utilizados en este campo mediante una aplicación específica sobre la base del desarrollo de una librería que utiliza algunas librerías estándar para el control de calidad que se encuentran en R. El nombre de la librería es qcr que corresponden a las siglas Quality Control and Reability.

La aplicación propuesta permite interactuar con las salidas gráficas en un entorno amigable para el mundo empresarial y académico evitando las restricciones de los paquetes actuales. Sin embargo la manipulación de los datos y su posterior análisis desde el entorno de R.

En el primer capítulo se presentan las definiciones necesarias para tener un marco conceptual de referencia dentro del conexto del control de calidad, además se realiza una revisión haciendo énfasis en las librerías propias de R para el control de calidad.

Para este trabajo, entre los diferentes paquetes de R que existen para la implementación, se ha elegido el paquete shiny desarrollado por el grupo Rstudio puesto que es muy versátil, fácil de implementar y permite el desarrollo de aplicaciones en ambiente web.

En el capítulo 2, se describen los gráficos de control de calidad por variable (\bar{x} , \bar{x} .one, R y S), por atributo (p, u y c) y por dependencia (CUSUM y EWMA) para el caso univariante y en el caso multivariante se tienen los gráficos T^2 , MCUSUM y EWMA. Todos estos gráficos han sido implementados en la librería qcr elaborada en este trabajo en base a los paquetes qcc (quality control charts) y MSQC (Multivariate Statistical Quality Control).

En el tercer capítulo, se realizan algunos ejemplos del uso de la librería y las salidas gráficas de la aplicación, adicionalmente en el anexo se describen las funciones desarrolladas en el paquete qcr. Finalmente en el capítulo cuatro se presentan la conclusiones y las futuras líneas de desarrollo.

En esta primera versión del paquete qcr (Quality Control and Reability) solo se han implementado herramientas para el control de la calidad para un control durante el proceso y no se ha incluido ninguna función para el control en el diseño (Reability). Sin embargo, se ha diseñado el paquete con un esquema modular de tal manera que extender su funcionalidad sea muy fácil y no dependa de las funciones anteriores desarrolladas.

Agradecimientos

A mis directores Dr. Salvador Naya Fernández y Dr. Rubén Fernández por brindarme su amistad y compartir sus conocimientos durante el desarrollo de este trabajo.

A todos los Profesores del Máster por compartir sus conocimientos científicos y experiencias. En especial a Ricardo Cao, Germán Aneiros, Jose Antonio Vilar, Wenceslao González Mateiga, Alberto Casal, Rosa Crujeiras y Juan Carlos Pardo.

Dedicatoria

A mi querida familia

Índice general

Resumen	2
Agradecimientos	4
Dedicatoria	5
1. Introducción	8
1.1. Antecedentes	8
1.2. Definiciones	9
1.3. Software R	10
1.3.1. Paquetes para elaborar gráficos de control de calidad	11
1.3.2. Paquete Shiny (Web Application Framework for R)	13
2. Consideraciones básicas	15
2.1. Breve reseña histórica de los gráficos de control	15
2.2. Generalidades	17
2.3. Bases estadísticas para la construcción de los gráficos de control	19
2.3.1. Gráficos univariante	19
2.3.2. Otros gráficos univariantes	20
2.3.3. Gráficos multivariantes	21
3. Gráficos de Control de Calidad (paquete - qcr)	25
3.1. Gráficos de control univariante	25
3.1.1. El Control de fabricación por variables	25
3.1.2. El Control de fabricación por atributos	46
3.1.3. El Control de fabricación por dependencia	58
3.2. Gráficos de control multivariante	64
3.2.1. Control de proceso multivariantes basados en el supuesto de normalidad	64

3.3. Interfaz gráfica para el control de calidad de procesos industriales . . .	75
4. Conclusiones y futuras líneas de desarrollo	80
A. Interfaz gráfica (Módulos)	82
Bibliografía	85

Capítulo 1

Introducción

1.1. Antecedentes

Se puede asumir que a lo largo de toda la historia siempre ha existido un interés por controlar y mejorar la calidad. Este interés está íntimamente ligado a la comprensión de la relación existente entre la calidad y la 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 concepto de control de calidad se ha extendido desde las primeras definiciones basadas en la idea de ajustar la producción a un modelo estándar, a la satisfacción de los requerimientos del cliente y englobar a todos los participantes, ampliando también la aplicación no sólo a productos manufacturados sino a todos los procesos industriales o al sector de los servicios.

Dependiendo del instante (antes, durante o después) de control, se puede controlar la calidad de un proceso industrial o servicio mediante tres enfoques, sin embargo en este trabajo sólo se considera la implementación de herramientas estadísticas para el control de calidad para el enfoque que corresponde al control durante el proceso industrial.

El enfoque cuando el control se realiza antes de la fabricación en el contexto industrial se conoce como calidad de diseño; el Control Estadístico de Procesos es utilizado durante la fabricación y se lo conoce como control preventivo; y finalmente cuando el control se lo realiza después de la fabricación se utiliza el muestreo de aceptación de lotes y se lo conoce como control de recepción.

En todos los enfoques del control de calidad se utiliza el método científico para la extracción eficiente de la información potencial generada por todo el proceso,

con el fin de una toma racional de decisiones que permita minimizar la variabilidad presente en el mismo. Siendo el recurso utilizado 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 trabajo se implementa una aplicación en la que se desarrollan la mayoría de los gráficos de control de calidad que son la principal herramienta estadística utilizada en el Control Estadístico de Proceso (control preventivo). Cabe mencionar que la aplicación es modular, es decir, en un futuro se pueden ir adaptando más herramientas que se utilizan en los diferentes enfoques.

1.2. Definiciones

Con el propósito de tener un marco conceptual básico sobre la terminología utilizada en este trabajo y en general en el ambiente empresarial se presentan a continuación algunas definiciones desde una perspectiva un poco más industrial.

La **competitividad** se define como la capacidad de optimizar la satisfacción de los clientes considerando el menor precio y fijado a un nivel de calidad. El concepto de competitividad está íntimamente relacionado con la calidad del producto y el nivel de precios.

Por otro lado, se entiende por **productividad** la capacidad de aumentar la producción de un producto fijando un nivel de calidad con los mismos o menos recursos. Por tanto, el precio y la calidad están ligados a la productividad que inciden directamente en la competitividad de una empresa ya que al aumentar la calidad en los productos ya no se tienen que reprocesar productos defectuosos y con esto se eliminan el desperdicio y el tiempo de la mala calidad y por tanto los costos de producción disminuyen logrando con esto disminuir los precios.

Se pueden encontrar diferentes definiciones en función de la fuente o de la época cuando se pretende aclarar que se entiende por calidad. Una de las más aceptadas es la definición que aporta la conocida norma ISO, que define la **calidad de un producto** como el "Conjunto de propiedades y características de un producto o servicio que le confieren su aptitud para satisfacer unas necesidades explícitas o implícitas".

Siempre existirá una variabilidad inherente durante el proceso de fabricación y por tanto se parte de la base que el conjunto de propiedades y características que se las considerará como características medibles de calidad de un producto o servicio

presentan cierta variabilidad. Esta variabilidad, lógicamente, será indeseable y un objetivo importante será analizarla y reducirla o, al menos, mantenerla dentro de unos límites de especificación.

Para controlar e intentar medir, y si es posible reducir, esa variabilidad inherente al producto o servicio se utilizan las técnicas de **control de la calidad**, que pueden definirse, según la norma ISO, como el "Conjunto de técnicas y actividades de carácter operativo, utilizadas para verificar los requerimientos relativos a la calidad del producto o servicio".

Dentro del conjunto de técnicas se tienen las herramientas aportadas por la Estadística. Específicamente el conjunto de técnicas estadísticas utilizadas para medir la calidad y compararla con las especificaciones o requisitos del cliente, con la intención de poder tomar medidas correctivas para su mejora, que en la mayoría de los casos estarán orientadas a la reducción de la variabilidad y que se define como **control estadístico de la calidad**. Es muy usual denominar al control estadístico de la calidad control estadístico de procesos (CEP) sobre todo desde la perspectiva industrial

Una de las principales herramientas utilizadas en el control estadístico de procesos son los **gráficos de control de calidad** que son gráficos con una línea límite superior y una línea límite inferior donde se representan los valores de alguna variable estadística para una serie de muestras sucesivas. El gráfico incluye también una línea central que corresponde al valor medio de las observaciones.

Como se mencionó anteriormente existe una variabilidad inherente al proceso de fabricación que se la conoce como **causas no asignables (comunes o aleatorias)**, por otra parte cuando la variabilidad aumenta por efectos predecibles que normalmente son muy pocos se conoce como **causas asignables**. Con respecto al estado de un proceso, se dice que un **proceso esta bajo control** cuando sólo existen causas no asignables.

1.3. Software R

En esta sección se presenta una visión general sobre el lenguaje R que es el motor de desarrollo de la aplicación computacional implementada en este trabajo y se menciona los paquetes para el control de calidad que hasta la fecha (última actualización mayo 2013), se encuentran alojados en la página web <http://www.r-project.org/>. Además se explican los criterios considerados para desarrollar un paquete para la implementación de los gráficos de control de calidad y la utilización del paquete

Shiny para el desarrollo de la interfaz gráfica.

R es un lenguaje y entorno de programación para análisis estadístico. Se trata de un proyecto de software libre cuyos primeros pasos se deben a Ross Ihaka y Robert Gentleman del Departamento de Estadística de la Universidad de Auckland, aunque las sucesivas versiones son controladas y desarrolladas por el R Development Core Team del que forman parte muchos colaboradores de todo el mundo. R se distribuye gratuitamente bajo los términos de la licencia pública general (GNU) y compila y se ejecuta en una diversidad de plataformas: UNIX, Windows y MacOS.

El lenguaje de programación incluye condicionales, lazos, y funciones. Con respecto a los argumentos de las funciones que están implementadas o las que el usuario crea pueden ser objetos de R, datos, fórmulas, expresiones, etc. En el entorno de R se utilizan paquetes que son un conjunto de funciones sobre una misma temática que proporcionan las facilidades para manejar conjuntos de funciones o datos. Cada función perteneciente a un paquete viene con la documentación correspondiente. Por ejemplo, el paquete denominado **base** constituye el núcleo de R y contiene las funciones básicas del lenguaje para leer y manipular datos, algunas funciones gráficas y algunas funciones estadísticas.

1.3.1. Paquetes para elaborar gráficos de control de calidad

Hasta el mes de Julio de 2013 en R se pueden utilizar para el control de calidad de procesos industriales dos paquetes (**qcc** y **IQCC**) que tienen implementado gráficos de control univariantes y/o multivariantes y tres paquetes (**MCUSUM**, **MEWNA** y **MSQC**) que se centran en el caso multivariado y sólo un paquete (**mnsqc**) para el caso noparamétrico. Cabe mencionar que existen otros paquetes para el control de calidad que se enfocan en la implementación de otras herramientas estadísticas tales como **qualityTools** y **SixSigma** especialmente orientados a la metodología Seis Sigma, aunque no tienen gráficos de control específicos implementan los del paquete **qcc**.

En el Capítulo 2 se describen los métodos de construcción de los diferentes gráficos de control implementados en estos paquetes (**xbar**, **xbar-one**, **S**, **R**, **p**, **np**, **c**, **u**, **ewna**, **cusum**, **mewna**, **mcusum** y T^2).

El paquete **qcc** ha sido desarrollado por el profesor Luca Scrucca del Departamento de Economía, Finanzas y Estadística de la Universidad de Perugia y como se puede observar en la Tabla 1.1 se presentan las implementaciones de los gráficos Shewhart de control de calidad por variable y atributo; así también los gráficos Cu-

Paquete	Variable	Atributo	Dependencia	Multivariante
qcc	xbar, xbar-one, S y R	p, np, c, y u	cusum y ewma	T^2
IQCC	xbar, S y R	p, c y u	-	T^2

Cuadro 1.1: Tipos de Gráficos de Control de Calidad implementados en los paquetes **qcc** (Quality Control Charts) y **IQCC** (Improved Quality Control Charts)

sum y Ewma para variables continuas con dependencia; y para el caso de los Gráficos de control multivariable tiene implementado T^2 Hotelling. Adicionalmente tiene implementado funciones para evaluar y calcular las características de funcionamiento de las curvas de operación OC, análisis de capacidad de proceso. Además se pueden realizar los diagramas de Pareto y de causa y efecto.

Por otro parte, el paquete **IQCC** es mantenido por el profesor Emanuel P. Barbosa del Instituto de Matemáticas de la Universidad Estatal de Campinas y como se puede observar de la Tabla 1.1 tiene implementados un número menor de gráficos de control.

Los paquetes **MCUSUM** (Multivariate Cumulative Sum Control Chart), **MEWMA** (Multivariate Exponentially Weighted Moving Average Control Chart) y **MSQC** (Multivariate Statistical Quality Control) han sido desarrollados por Edgar Santos Fernández de la empresa de Telecomunicaciones de Cuba.

El paquete **MSQC**, es un conjunto de herramientas para el proceso de control multivariante. Contiene las principales alternativas de gráficos de control multivariante tales como: Hotelling (T^2), Chi cuadrado, MEWMA, MCUSUM y gráfico de control de la Varianza Generalizada. Además, incluye algunas herramientas para evaluar la normalidad multivariada.

En el caso en que no se pueda garantizar la normalidad de los datos, se podría utilizar gráficos de tipo no paramétrico que no requieren esta hipótesis, se tiene el paquete **mnspec** (Multivariate Nonparametric Statistical Process Control) desarrollado por Martín Bezener y Peihua Qiu de la Universidad de Minnesota. En este paquete se tiene una versión del procedimiento CUSUM. En el paquete **qcr** a un futuro muy cercano se implementará gráficos no paramétricos que utiliza el concepto de profundidad como es el caso de los gráficos de control de calidad (r, Q, D y S) presentados por Regina Liu (1990).

Las principales razones de desarrollar el paquete **qcr** en este trabajo son:

- Tener en un solo paquete integrados los gráficos de control de calidad más

utilizados para el caso paramétrico y noparamétrico.

- Mejorar las salidas gráficas de los paquete de control de calidad antes mencionados.
- Diseñar las funciones de tal forma que la interacción con el desarrollo de la interfaz gráfica sea más eficaz ya que la filosofía de desarrollo esta en base al paquete **Shiny**.
- Facilitar la inclusión de nuevas funciones para la creación de gráficos de control ya que el diseño es modular, es decir, no depende de las anteriores implementación.
- No redundar en código ya que se utiliza las funciones implementadas en los paquetes qcc y MSQC.
- Permitir utilizar las funciones de otros paquetes en el caso de que ya estén implementadas para el cálculo de las medidas estadística y así solo se preocupa de las salidas gráficas.
- Implementar gráficos de control de calidad multivariante noparamétricas que no se encuentran actualmente desarrolladas en ningún paquete.

1.3.2. Paquete Shiny (Web Application Framework for R)

El criterio utilizado para la búsqueda del paquete ha utilizar para el desarrollo de la aplicación se centró en encontrar un paquete que combine las fucionalidades del paquete qcr desarrollado en este trabajo con la simplicidad de una página web y la mejora de las salidas gráficas. Web Application Framework for R

En el CRAN de R se pueden encontrar varios paquetes de tipo **GUI** (Interfaces Gráficas de Usuario) que se podría utilizar para implementar la aplicación para el control de calidad, sin embargo la mayoría están sin mantenimiento o bien es demasiado difícil utilizarlos para la implementación de una aplicación.

Se ha considerado el paquete **Shiny** ya que permite que sea muy simple para los usuarios de R diseñar aplicaciones web interactivas que seán fáciles de usar. Intendiendo por interactivas que deje que los usuarios elijan los parámetros de entrada usando los controles amigables como deslizadores, menús desplegable y campos de texto. Además permite de manera muy fácil implementar cualquier número de salidas como gráficos, tablas y resúmenes.

El paquete **Shiny** ha sido desarrollado por el grupo RStudio y fue lanzado a finales del 2012 con gran éxito entre los usuarios de R y que cuenta con actualizaciones con mucha frecuencia, actualmente se tiene la versión 0.5.0.

Capítulo 2

Consideraciones básicas

El objeto de este capítulo es presentar los inicios de los gráficos de control como una herramienta del control estadístico de procesos, empezando por los gráficos propuestos por Shewhart y la importante aportación realizada por Hotelling. Así también presentar algunos aspectos generales y las bases estadísticas (distribuciones de probabilidad) de los gráficos de control implementados en el paquete *qcr*.

2.1. Breve reseña histórica de los gráficos de control

El control estadístico de procesos se fundamenta en la propuesta realizada por el Dr. Walter Shewhart que los empleó por vez primera cuando trabajaba en la empresa Bell Telephone, en la década de los 20 del siglo pasado. El Dr. Shewhart mediante los gráficos de control los utilizaba para identificar los problemas ocurridos durante el proceso que pueden ser por causas asignables o causas no asignables; su identificación daba como resultados la reducción de la variabilidad y la mejora del proceso.

El Dr. Shewart indicaba que el control de procesos consta de dos fases. La fase de desarrollo, que ayuda a conocer el comportamiento estadístico del proceso, y permite determinar los límites de control para el estimador del parámetro analizado para esto se tiene que eliminar causas asignables si existieren. Por otro lado, la fase de madurez, en la cual se mide la capacidad del proceso para cumplir con los estándares de calidad inicialmente propuestos; se identifica el promedio del número de muestras antes de obtener falsas alarmas; y ayuda a reducir el tiempo o el número de muestras necesarias para detectar pequeños cambios en el proceso.

Los primeros gráficos de control propuestos por Shewhart, fueron para monito-

rizar el promedio (\bar{x}) y la variación (R o S), estos gráficos de control se aplican a variables cuantitativas. Posteriormente aparecieron los gráficos c y p para monitorizar variables cualitativas tales como el número de defectos por lotes y la proporción de defectos por lotes, respectivamente. Este tipo de gráficos pertenecen al control univariante para variables cuantitativas y cualitativas pero no consideran el efecto de dependencia en el tiempo.

Para el caso cualitativo pero considerando el componente de dependencia en el tiempo se incorporarán dos nuevos gráficos para detectar pequeñas variaciones en el proceso, uno introducido por E. S. Page (1954) denominado gráfico de sumas acumuladas (CUSUM) y el segundo introducido por S. W. Roberts (1959) denominado gráfico de promedios móviles exponencialmente ponderados (EWMA). Estos gráficos van acumulando información obtenida de las muestras anteriores y de esta manera pueden detectar de forma más rápida pequeños cambios en el proceso.

Actualmente es frecuente que el control estadístico de procesos sea de tipo multivariante, es decir, se monitorizan diversas variables relacionadas a la calidad. La principal aportación en este campo fue la realizada por el Profesor Harold Hotelling (1947), quien propuso un gráfico de control multivariante basado en la distancia de Mahalanobis (1936), utilizando el vector de promedios y la matriz de covarianzas de una variable aleatoria con distribución normal multivariada.

El gráfico de control multivariante, denominado gráfico de control T^2 de Hotelling, el cual monitoriza la distancia entre el vector de promedios esperado y el vector de promedios observados, considerando su matriz de covarianzas y que utiliza un límite de control (CL) el cual indica si la distancia entre los dos vectores es lo suficientemente grande para declarar al proceso como fuera de control.

Posteriormente en este mismo campo surgen aportaciones tales como la realizada por Lowry (1992) a través de la extensión del gráfico de control EWMA para el caso multivariante MEWMA, y la de Crosier en 1988 había presentado una versión multivariada del gráfico de control de sumas acumuladas MCUSUM. Los gráficos multivariantes mencionados están sujetos al supuesto de normalidad multivariada que podría ser una limitación para los casos donde la hipótesis sobre la distribución impuesta para estos modelos paramétricos multivariantes no se cumpla. En este sentido se presentan las aportaciones de Regina Liu que ha diseñado los denominados gráficos DDChart, que son gráficos de control de tipo no paramétrico, basados en el concepto de profundidad de datos.

2.2. Generalidades

Como ya se comentó en la sección anterior, dentro del control estadístico de procesos la herramienta clave son los gráficos de control, en los que se representa la evolución del valor de un estadístico calculado sobre las mediciones de una característica de calidad efectuadas sobre los elementos de muestras sucesivas. Con los gráficos de control se detecta la aparición de salidas de control o cambios importantes en el proceso, es decir, visualizar cuando se presentan causas no asignables (aleatorias) de cuando las causas son asignables.

En el gráfico se van representando de forma correlativa los valores muestrales con objeto de poder monitorizar el proceso. Se acostumbra a unir los valores representados mediante líneas poligonales para poder visualizar con mayor facilidad la salida de los límites de control, o la presencia anómala de secuencias no aleatorias (rachas). Debe mencionarse que aún cuando todos los valores muestrales se encuentren en los límites de control pueden presentarse secuencias de valores por encima o por debajo de la línea central que indiquen claramente una alteración en el proceso, para ello serán útiles realizar distintos contrastes que permitan detectar estas salidas de control.

En general para interpretar los patrones de un gráfico de control existen las siguientes pautas:

- Tendencias o rachas: Si el desplazamiento de la media o de la variabilidad es paulatino a lo largo del tiempo.
- Periodicidades: Son repeticiones cíclicas de alguna de las gráficas de control.
- Inestabilidad: Consiste en la presencia de grandes fluctuaciones que, en ocasiones, pueden provocar que algún valor caiga fuera de los límites de control.
- Sobreestabilidad: Es el fenómeno contrario y se da cuando la variabilidad es menor de la esperada.

De manera más específica se cuentan con diferentes criterios para analizar los gráficos de control y tomar decisiones, por ejemplo, se tiene la reglas de la Western Electric. Estas reglas dicen que una salida de control estaría indicada si ocurre uno de los siguientes casos:

- Un punto cae fuera de los límites 3 sigma.

- Dos de tres puntos caen fuera de los límites dos sigma.
- Cuatro de cinco puntos consecutivos se encuentran a una distancia de un sigma o más de la línea central.
- Ocho puntos consecutivos se encuentran al mismo lado de la línea central.

Saber si el estadístico calculado a partir de la i -ésima muestra está o no dentro de los límites de control, puede ser definido como un contraste de hipótesis estadística; para ello partimos que la hipótesis nula H_0 debe ser la hipótesis que el experimentador asume como correcta. En el caso del control de calidad la hipótesis nula sería suponer que el proceso está bajo control en cada una de las muestra seleccionadas. Rechazar la hipótesis nula implica asumir como correcta una hipótesis complementaria, que en el contexto de los contrastes se denomina hipótesis alternativa y que suele denotarse por H_1 . Un punto que esté fuera de los límites de control puede interpretarse como un incumplimiento de la hipótesis nula y una aceptación de la alternativa, es decir, dar por contrastado que el proceso ha salido de control.

En un contraste de hipótesis, a la decisión de rechazar la hipótesis nula cuando ésta es cierta se la denomina error de tipo I (α), en este contexto de control es más habitual denominarlo o riesgo del vendedor, mientras que al error cometido al no rechazar la hipótesis nula cuando ésta es falsa se la denomina riesgo del comprador (error de tipo II (β)).

Por tanto, durante la monitorización de un proceso puede existir la presencia de falsas alarmas (error de tipo I), es decir, se puede tener un punto fuera de los límites de control cuando el proceso esta bajo de control. Esto genera la necesidad de calcular la longitud promedio de rachas o números de muestras a tomar, hasta que aparece una señal fuera de control o hasta que caiga un punto por primera vez fuera de los límites de control (ARL) para evaluar la efectividad del gráfico de control. Ya que cada toma de muestra cumple con las características de un ensayo Bernoulli, entonces, la variable “número de muestras a tomar hasta que por primera vez una muestra se sitúe fuera de los límites de control”, tiene una distribución binomial negativa con parámetros $r = 1$ y $p = 1 - \beta$, Carot (1998); entonces, $ARL = \frac{1}{(1-\beta)}$. Para el caso donde se desee calcular el ARL considerando patrones o rachas no comunes se debe recurrir al empleo de simulaciones o modelación mediante procesos estocásticos.

En los contrastes de hipótesis se mide la potencia del test mediante la denominada curva de potencia, sin embargo, en el caso de los gráficos de control se utiliza la inversa

de esta curva de potencia que se denomina curva característica de operación (curva OC de las siglas Operating Characteristics) y que representa visualizar en un gráfico el error de tipo II (también llamado riesgo β) en función de distintos valores para la hipótesis alternativa. Es frecuente que dicha curva se construya representando este error en función de la magnitud del cambio que se pretende estudiar (normalmente expresado en unidades de desviación estandar).

Con respecto a la selección del tamaño de la muestra suele hacerse uso de las curvas OC. Si se representan para distintos tamaño de muestra estas curvas puede ser un instrumento eficaz para analizar el poder del gráfico; sin embargo, en el caso de controlar una variable con poca dispersión suelen emplearse muestras pequeñas de tamaño 4 o 5 a intervalos cortos de tiempo. Cuando se estudia el control por atributos las muestras necesitarán ser de mayor tamaño con objeto de poder aplicar las correspondientes aproximaciones de las distribuciones discretas por una continua, como por ejemplo en el caso de usar la distribución normal como límite de la binomial.

Por otra parte la frecuencia de muestreo es otro factor importante a tener en cuenta, lo habitual es considerar cuestiones económicas y del propio proceso y encontrar un equilibrio entre el tamaño de las muestras y la frecuencia.

2.3. Bases estadísticas para la construcción de los gráficos de control

2.3.1. Gráficos univariante

En el control estadístico de procesos, para el caso univariante, cuando la característica de interés es medible, generalmente sigue una distribución de probabilidad normal $X \in N(\mu, \sigma)$; si no lo tuviera, se puede hacer uso del Teorema Central del Límite o de transformaciones matemáticas para aproximar o asumir su normalidad.

Para características medibles, existe interés principalmente por estimar:

- la media μ del proceso a través de \bar{X} . Donde $\bar{X} = \sum_{i=1}^n \frac{x_i}{n}$; y n es el tamaño de la muestra. Con X_i normalmente distribuida se tiene que $\bar{X} \in N(\mu, \frac{\sigma}{\sqrt{n}})$ y cuando X no proviene de una distribución Normal, la aplicación del teorema central del límite soluciona el problema.
- la variabilidad a través del rango R en cada muestra. Dado que $R = \max(x_i) - \min(x_i)$ y aplicando el Teorema de Crámer, el cual indica que si X e Y son

variables aleatorias independientes y ambas están normalmente distribuidas y son sus respectivas medias (μ_X, μ_Y) y varianzas (σ_X, σ_Y) entonces $V = X - Y \in N(\mu_X - \mu_Y, \sigma_X + \sigma_Y)$ de tal forma es fácil observar que R tiene distribución normal.

- la desviación estándar σ , a través de S o σ^2 utilizando S^2 . Donde $S^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$, y por tanto $S = \sqrt{S^2}$.

Por otro lado, si las características de calidad es una cualidad (defectuoso o no defectuoso) pero puede ser cuantificada, se tienen los siguientes estimadores:

Para características cuantificables, los estimadores utilizados son:

- número de elementos D que no cumplen con las características de calidad (disconformes) y que pertenecen a una muestra de tamaño n . Su función de distribución de probabilidad es binomial, donde $E(D) = np$, y $Var(D) = np(1 - p)$.
- fracción de disconformes p , siendo su estimador $\hat{p} = \frac{D}{n}$. Haciendo uso de esta última relación, aparece el gráfico de control para el número de disconformes np , siendo fácil deducir que $D = n\hat{p}$. Tiene función de distribución de probabilidad binomial; $E(\hat{p}) = \frac{E(D)}{n} = p$, y $Var(\hat{p}) = \frac{Var(D)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$.
- cantidad de defectos c o disconformidades por unidad de observación. Tiene una función de distribución de probabilidad Poisson, donde la $E(X) = Var(X) = c$. Además $\hat{c} = \bar{c}$.
- promedio de defectos por unidades de inspección u , cuando $u = \frac{x}{n}$, siendo x el número de disconformidades totales en una muestra de n unidades de inspección. Siendo $X \in Poisson(\lambda)$ se tiene $E(u) = E(\frac{X}{n}) = \frac{E(X)}{n} = \frac{\lambda}{n}$ y $Var(u) = Var(\frac{X}{n}) = \frac{Var(X)}{n^2} = \frac{\lambda}{n^2}$. Haciendo el cambio de variable $p = \frac{\lambda}{n}$ y utilizando la aproximación de la Poisson a la Normal, se tiene que $u \in N(p, \frac{p}{n})$. Si no se cuenta con el parámetro p , se lo estima a través de $\hat{p} = \bar{u}$ con $u \neq \bar{u}$.

2.3.2. Otros gráficos univariantes

En este trabajo se implementan otros gráficos para el caso univariante además de los propuestos por Walter Shewhart para el control de procesos donde no sólo usán información contenida en la última muestra para la toma de decisiones si no

que además incluyen toda la información de los puntos anteriores. Esta información acumulada es de gran importancia en lo que podríamos denominar la fase de monotorización.

Los gráficos de sumas acumuladas o gráficos CUSUM y los gráficos con medias ponderadas exponencialmente o gráficos EWMA no se basan en las observaciones individuales, o promedios de una muestra de ellas, sino en la acumulación de la información. Por esta razón se les denomina también gráficos con memoria.

El gráfico de sumas acumulativas o gráficos CUSUM (Cumulative Sum) que basa en el análisis de toda o la mayor parte de la muestra. Para este fin se usa el estadístico s_r :

$$s_r = \sum_{i=1}^r (\bar{x}_i - \mu_0)$$

Este estadístico tiene en cuenta las r primeras submuestras para tomar la decisión sobre la muestra r ésima. Considerando todas las diferencias de cada media muestral (\bar{x}_i) con la media del proceso (μ_0). Si se llama a estas diferencias $d_i = \bar{x}_i - \mu_0$. La idea del gráfico CUSUM es representar las sumas acumuladas de estas diferencias, es decir los valores: $s_1 = d_1, s_2 = d_1 + d_2, \dots$, de esta forma en cada instante r se considera también la información histórica de los valores anteriores. Un desajuste pequeño se irá así acumulando haciendo más fácil su detección. Bajo el supuesto de normalidad la distribución del estadístico $s_r \in N\left(r(\mu - \mu_0), \frac{r\sigma_0^2}{n}\right)$.

Para los gráficos EWMA El estadístico dibujado es la media ponderada de la observación actual y todas las observaciones precedentes, donde la media previa recibe más peso, es decir, $Z_t = \lambda x_t + (1 - \lambda) Z_{t-1}$, $0 < \lambda < 1$.

Donde: $Z_0 = \mu$, Z_t y Z_{t-1} son las medias móviles ponderadas exponencialmente en el tiempo actual t y $t - 1$ respectivamente, x_t es la observación presente y λ es el factor de ponderación para la observación actual.

Si se cumple la hipótesis de independencia y la estabilidad sobre los x_t , se tiene que $E(Z_t) = \mu$ y $Var(Z_t) = \sigma^2 \frac{\lambda(1-\lambda)^{2t}}{2-\lambda}$

2.3.3. Gráficos multivariantes

El interes en el control estadístico del proceso multivariante está en controlar un vector de p componentes, que desde el punto de vista estadístico se denominaría variable aleatoria p -dimensional, o vector aleatorio, y se podría denotar por $\vec{X} = (X_1, X_2, \dots, X_p)$, y sería un conjunto de p variables aleatorias unidimensionales

definidas sobre el mismo espacio de probabilidad. Estas variables podrán medir datos de tipo discreto o de tipo continuo, e incluso con frecuencia tener unas componentes discretas y otras continuas.

La esperanza del vector aleatorio \vec{X} , también llamada vector de medias, es el vector formado por las medias individuales:

$$E(\vec{X}) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \dots \\ E(X_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_p \end{pmatrix} = \vec{\mu}$$

La matriz de varianzas-covarianzas asociada al vector de medias, se define por

$$Cov(\vec{X}) = \Sigma = (\sigma_{ij}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \dots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{pmatrix} = E\left(\left(\vec{X} - \vec{\mu}\right)\left(\vec{X} - \vec{\mu}\right)^t\right)$$

donde $\sigma_{ij} = Cov(X_i, X_j)$

En particular, la matriz Σ es simétrica, en su diagonal aparecen las varianzas de las variables y fuera de ella las covarianzas entre cada par de variables. También puede comprobarse que la matriz de varianzas-covarianzas es semidefinida positiva. El determinante de Σ que coincide con el producto de todos sus autovalores, que son positivos es la llamada varianza generalizada de la variable multidimensional, pues conserva algunas de las propiedades de la varianza de una sola variable y mide la dispersión conjunta del vector de variables aleatorias.

Una medida adimensional de la relación entre las variables componentes del vector aleatorio es la matriz de correlaciones, que es la formada por los coeficientes de correlación lineal, $\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$, siendo $\sigma_i^2 = \sigma_{ii}$. Esto da lugar a una matriz $P_{p \times p} = (\rho_{ij})$, con unos en la diagonal y elementos comprendidos entre 1 y -1 fuera de ella (los coeficientes de correlación de unas variables con otras).

A la hora de estimar cantidades como las anteriormente mencionadas (el vector de medias, la matriz de varianzas-covarianzas o la de correlaciones, por ejemplo) se puede proceder de forma análoga a lo ya conocido para variables aleatorias unidimensionales. Supongamos que se dispone de una muestra de datos x_{ijk} correspondientes a la observación i -ésima ($i = 1, 2, \dots, n$) de la j -ésima característica de calidad ($j = 1, 2, 3, \dots, p$) en la k -ésima muestra ($k=1, 2, \dots, m$), que pueden estructurarse mediante la matriz. El vector de medias poblacional, $\vec{\mu}$, puede estimarse mediante el llamado vector de medias muestral dado por:

$$\bar{x}_{ik} = \frac{1}{n} \sum_{j=1}^n x_{ijk}$$

y la estimación de la varianza para cada elemento del vector aleatorio viene dada por:

$$s_{jk}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ijk} - \bar{x}_{jk})^2$$

La matriz de varianzas-covarianzas poblacional puede estimarse también mediante su análoga muestral al igual que el vector de medias. Así dadas las covarianzas entre la característica de calidad j ésima y la característica de calidad h ésima para la k ésima muestra podrá estimarse mediante la expresión:

$$s_{jhk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ijk} - \bar{x}_{jk})(x_{ihk} - \bar{x}_{hk}), \quad k = 1, 2, \dots, m; j \neq k$$

En base a los anteriores estimadores se definen los siguientes estimadores:

$$\bar{\bar{x}}_j = \frac{1}{m} \sum_{k=1}^m \bar{x}_{jk}, \quad j = 1, 2, \dots, p$$

$$\bar{\bar{s}}_j^2 = \frac{1}{m} \sum_{k=1}^m s_{jk}^2, \quad j = 1, 2, \dots, p$$

$$\bar{s}_{jh} = \frac{1}{m} \sum_{k=1}^m s_{jhk}, \quad j \neq h$$

Estos cálculos pueden reescribirse en forma matricial:

$$\bar{\bar{x}} = \begin{pmatrix} \bar{\bar{x}}_1 \\ \bar{\bar{x}}_2 \\ \dots \\ \bar{\bar{x}}_p \end{pmatrix}$$

$$S = \begin{pmatrix} \bar{\bar{s}}_1^2 & \bar{s}_{12} & \dots & \bar{s}_{1p} \\ & \bar{\bar{s}}_2^2 & \dots & \bar{s}_{2p} \\ & & \ddots & \vdots \\ & & & \bar{\bar{s}}_p^2 \end{pmatrix}_{p \times p}$$

Estas estimaciones del vector de medias y de las matrices de varianzas-covarianzas son el punto de partida de muchas de las técnicas del control estadístico multivariante. Por ejemplo, se necesitan estas estimaciones para aplicar la distribución multiva-

riante más utilizada (como ya ocurría también en el caso univariante) es la normal multidimensional para el control estadístico de procesos multivariantes, que se denota por $N_p(\vec{\mu}, \Sigma)$, cuya densidad viene dada por

$$f(\vec{x}) = (2\pi)^{-\frac{p}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\vec{x} - \vec{\mu})^t \Sigma^{-1}(\vec{x} - \vec{\mu})\right]$$

,siendo $\vec{\mu}$ un vector de p componentes y Σ una matriz de dimensión $p \times p$ simétrica y definida positiva. Donde $\vec{\mu}$ resulta ser el vector medias de esta distribución y que Σ es la matriz de varianzas-covarianzas de la misma.

Los gráficos de control multivariantes sólo son válidos cuando la variable sigue una distribución normal multivariante, además la representación gráfica no es posible para $p > 3$ y para $d = 2$ el gráfico pierde el orden cronológico. Una forma de sobrellevar estas limitaciones es mediante el control no paramétrico de procesos multivariantes que se basa en el concepto de profundidad de datos.

Capítulo 3

Gráficos de Control de Calidad (paquete - qcr)

El objeto de este capítulo es revisar la metodología para la obtención de los gráficos de control del tipo univariado y los gráficos de control multivariados. Se presentan algunos ejemplos de cada tipo de gráfico de control y se utiliza el paquete desarrollado para la construcción de los gráficos de control. Para mayor información sobre las funcionalidades del paquete se anexa a este trabajo el tutorial del paquete.

3.1. Gráficos de control univariante

3.1.1. El Control de fabricación por variables

Para el control de una característica medible y representada por una variable aleatoria X , existen dos tipos de gráficos, por una parte gráficos que nos ayuden a controlar la posición (gráficos de la media o mediana) y por otra los que controlan la dispersión (gráficos de desviaciones, varianzas o recorridos).

La etapa inicial del procedimiento para el control consiste en elegir un número k de muestras de un determinado tamaño, donde se denotará por x_{ij} el j -ésimo elemento de la i -ésima muestra y n el tamaño de cada muestra. Para posteriormente, obtener un estimador de la media del proceso que en este caso sería la media muestral de las k muestras, que se denotará por $\bar{\bar{x}}$, y que corresponde con la media muestral conjunta de los nk datos:

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k x_i}{k} = \sum_{i=1}^k \sum_{j=1}^n \frac{x_{ij}}{N}$$

con $N = nk$

Para el caso de la estimación de la variabilidad:

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

Esta estimación de la varianza no tiene en cuenta el hecho de que los datos provienen de distintas submuestras y si estas difieren mucho de unas a otras, el valor de \hat{s}^2 será demasiado grande, subestimando la varianza del proceso. Es decir, con este método se combinaría la variabilidad dentro de las muestras con la variabilidad entre las muestras, pero lo ideal será que la variabilidad se base sólo en una medida que analice la situación dentro de las muestras, para lo que se utilizan estimaciones basadas en calcular la varianza o rango (recorrido) de cada submuestra por separado para luego utilizar su media.

En las siguientes secciones se presentan gráficos de control para monitorizar la posición (tendencia centralidad) y variabilidad del proceso, donde se considerarán mejores estimaciones de la varianza (variabilidad).

Para los ejemplos en esta sección se utiliza los datos “pistonrings (Piston rings data)” del paquete `qcc`, donde se mide el diámetro de los aros de los pistones fabricados. Los datos están en un data frame con 200 observaciones medidas en 25 muestras cada una de tamaño 5; para cada observación se tiene información de las las siguientes variables:

- diameter : a numeric vector
- sample : sample ID
- trial : trial sample indicator (TRUE/FALSE)

Gráficos de control (\bar{x}, R)

Uno de los gráficos de control de calidad para controlar la variabilidad del proceso es el basado en los rangos. El rango o recorrido de la muestra i ésima no es más que la diferencia entre la mayor y la menor de sus observaciones:

$$R_i = \max(x_i) - \min(x_i)$$

Bajo la hipótesis de que la distribución es normal, puede calcularse la distribución del rango muestral. En particular, su media es $d_2\sigma$, donde d_2 es una constante que

depende del tamaño muestral, n , y que suele darse tabulada pero al usar el paquete qcr se obtiene su valor automáticamente.

Para construir los gráficos de control para la media se puede realizar una estimación de la varianza a través del rango puesto que $\frac{R_i}{d_2}$ es un estimador insesgado de σ y, por tanto, su media (a lo largo de todas las muestras) también lo será.

Definiendo $\bar{R} = \frac{\sum_{i=1}^k R_i}{k}$, se tiene entonces que podemos estimar σ mediante $\frac{\bar{R}}{d_2}$ y construir los límites de control de los gráficos de medias (llamados ahora gráficos (\bar{x}, R)) mediante:

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

En donde $\bar{\bar{x}}$ es la media global de todos los datos o, lo que es equivalente, la media de las medias de cada submuestra: $\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} = \frac{\sum_{i=1}^k \sum_{j=1}^n x_{ij}}{N}$, con $N = kn$. Si todas las muestras provienen de la misma distribución entonces se tiene que $\frac{\sqrt{nd_2}(\bar{x}_i - \bar{\bar{x}})}{\bar{R}}$ tiene distribución aproximada $N(0, 1)$. Con lo cual cada valor \bar{x}_i debe estar en el intervalo $\left[\bar{\bar{x}} - \frac{3\bar{R}}{\sqrt{nd_2}}, \bar{\bar{x}} + \frac{3\bar{R}}{\sqrt{nd_2}}\right] = \left[\bar{\bar{x}} - A_2\bar{R}, \bar{\bar{x}} + A_2\bar{R}\right]$. Al igual que ocurría con d_2 la constante A_2 se calculan a través del paquete qcr. El paquete qcr no presenta los valores individuales para cada constante sino que se calculan directamente los siguientes límites y los representa gráficamente:

$$LCL = \bar{\bar{x}} - A_2\bar{R}$$

$$CL = \bar{\bar{x}}$$

$$UCL = \bar{\bar{x}} + A_2\bar{R}$$

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

También se puede tener otro gráfico de control mediante rangos donde se tiene una línea central con altura \bar{R} y líneas de control $D_3\bar{R}$ y $D_4\bar{R}$, estando D_3 y D_4 obtenidas por el paquete qcr. Para la determinación de D_3 y D_4 se debe de tener en

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

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

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

$$LCL = \bar{R} - 3\hat{\sigma}_R = \bar{R} - 3d_3 \frac{\bar{R}}{d_2} = D_3\bar{R}$$

$$CL = \bar{R}$$

$$UCL = \bar{R} + 3\hat{\sigma}_R = \bar{R} + 3d_3 \frac{\bar{R}}{d_2} = D_4\bar{R}$$

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

$$D_3 = 1 - 3 \frac{d_3}{d_2}$$

$$D_4 = 1 + 3 \frac{d_3}{d_2}$$

De manera análoga, el paquete qcr no presenta los valores individuales de las constantes D_3 y D_4 sino que se calculan de manera interna.

Ejemplo 1: Se realiza un gráfico de medias para monitorizar el diámetro de los piston rings.

- Primero se procede a cargar la librería qcr y los datos.

```
> library(qcr)
Loading required package: qcc
Loading required package: MASS
Package 'qcc', version 2.2
Type 'citation("qcc")' for citing this R package in publications.

Package qcr: Quality control and reliability
version 0.1-9 (built on 2013-06-22).
Copyright Miguel A. Flores Sanchez 2013.

> data(pistonrings)
```

```
> str(pistonrings)
'data.frame': 200 obs. of 3 variables:
 $ diameter: num 74 74 74 74 74 ...
 $ sample  : int 1 1 1 1 1 2 2 2 2 2 ...
 $ trial   : logi TRUE TRUE TRUE TRUE TRUE TRUE ...
```

- Se crea el objeto `qcd` (Quality Control Data) y el objeto `qcs` (Quality Control Statistics) del tipo `xbar`. Para indicar que el tipo de estimación de la varianza sea el visto en esta sección en el comando `qcs.xbar` se indica en la opción `std.dev = "UWAVE-R"` que es método de estimación por defecto.

```
> pistonrings.qcd <- qcd(pistonrings)
> class(pistonrings.qcd)
[1] "qcd"          "data.frame"
> res.qcs <- qcs.xbar(pistonrings.qcd)
```

- Con el comando `summary` se obtienen las estimaciones de la media y la varianza así como los límites de control y con el comando `plot` se realiza el correspondiente gráfico de control.

```
> summary(res.qcs)
xbar chart for pistonrings

Summary of group statistics:
      xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.02

Group sample size: 5
Number of groups: 40
Center of group statistics: 74.0036
Standard deviation: 0.01007094
```

Control limits:

```

      LCL      UCL
73.99009 74.01712

```

Beyond limits of control:

```
[1] 74.0196 74.0234
```

Violating runs:

```
[1] 74.0128
```

```
> plot(res.qcs, title = "Control Chart Xbar for pistonrings I")
```

Se puede observar tres puntos que indican que el proceso está fuera de control. Dos que salen fuera de los límites de control y un punto que está dentro de los límites pero que es el resultado de una racha. En la Figura 3.1 se tiene el correspondiente gráfico de control.

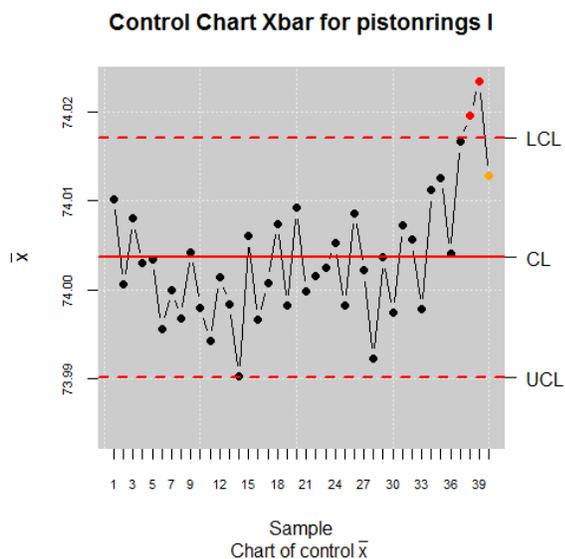


Figura 3.1: Gráfico de Control para las medias de la característica de calidad diámetro.

- Cómo se ha determinado en la etapa anterior el proceso no está bajo control y por tanto los límites de control no son los utilizados para efectuar la monitorización. Con la función `state.control` se eliminan los puntos fuera de control y devuelve un objeto del tipo `qcd`. Con este nuevo objeto se repiten las etapas

anteriores hasta que el proceso este bajo control, es decir, no existan puntos fuera de los límites o rachas.

```
> res.qcd <- state.control(res.qcs)
> res.qcs <- qcs.xbar(res.qcd)
> summary(res.qcs)
xbar chart for pistonrings
```

Summary of group statistics:

```
      xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.02
```

```
Group sample size: 5
Number of groups: 37
Center of group statistics: 74.00239
Standard deviation: 0.009992796
```

Control limits:

```
      LCL      UCL
73.98898 74.0158
```

Beyond limits of control:

```
[1] 74.0166
```

Number violating runs: 0

Aún existe un punto fuera de los límites de control (Figura 3.2), ejecutamos de nuevo el script.

```
> res.qcd <- state.control(res.qcs)
```

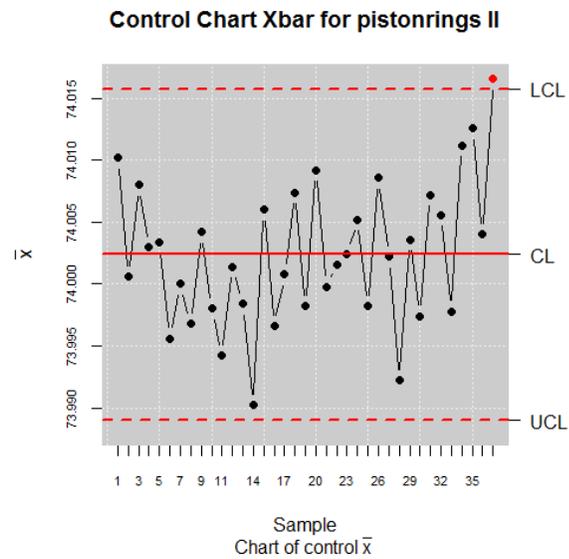


Figura 3.2: Gráfico de Control para las medias de la característica de calidad diámetro (un punto fuera de los límites de control).

```
> res.qcs <- qcs.xbar(res.qcd)
> summary(res.qcs)
xbar chart for pistonrings
```

Summary of group statistics:

```
      xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.01
```

```
Group sample size: 5
Number of groups: 36
Center of group statistics: 74.00199
Standard deviation: 0.01004347
```

Control limits:

```
      LCL      UCL
```

```
73.98852 74.01547
```

```
Number beyond limits: 0
```

```
Number violating runs: 0
```

Finalmente el proceso esta bajo control (ver Figura 3.3) y los límites ha utilizar para monitorizar el proceso son: (73,99, 74,01).

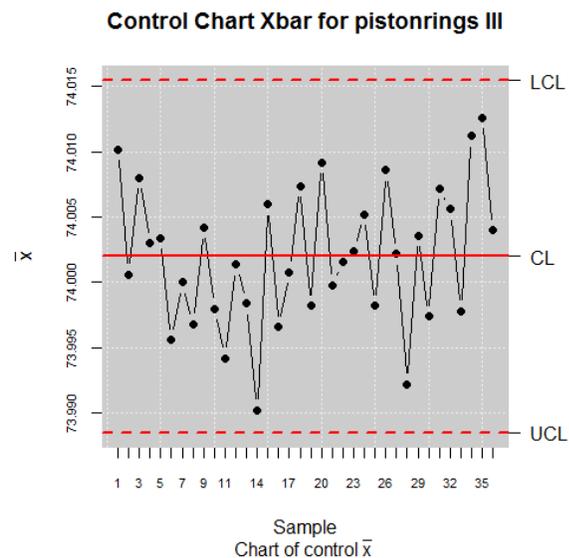


Figura 3.3: Gráfico de Control para las medias de la característica de calidad diámetro (proceso bajo control).

Ejemplo 2: En este ejemplo se parte de un proceso bajo control (Fase I) y una vez calculado los límites se realiza la monitorización del proceso (Fase II). Para la Fase I se considera las primeras 25 muestras de los datos de piston rings y para la Fase II se utilizan el resto de las muestras.

- Calculamos los límites para la Fase I.

```
> library(qcr)
> data(pistonrings)
> x <- droplevels(pistonrings[1:125,])
> y <- droplevels(pistonrings[126:200,])
```

```
> res.qcs <- qcs.xbar(x, data.name="Control Chart Xbar for pistonrings")
> summary(res.qcs)
xbar chart for Control Chart Xbar for pistonrings
```

Summary of group statistics:

```
      xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.00
Max.   :74.01
```

```
Group sample size: 5
Number of groups: 25
Center of group statistics: 74.00118
Standard deviation: 0.009785039
```

Control limits:

```
      LCL      UCL
73.98805 74.0143
```

```
Number beyond limits: 0
```

```
Number violationg runs: 0
```

```
> plot(res.qcs)
```

Los límites son (73,98805, 74,0143) y como se puede ver en la Figura 3.4 no se tiene puntos fuera de los límites de control.

- Para la Fase II se utiliza la función `merge.qcs` que permite crear un nuevo objeto `qcs` del tipo `xbar` que es la unión del objeto `qcs` creado a partir de las primeras 25 muestras y un data frame que almacena el resto de las muestras.

```
> res.qcs <- merge.qcs(x = res.qcs, y = y[,c(1,2)])
```

Warning message:

```
In qcc(data = x, sizes = sizes, type = type, confidence.level = conf.nsigma,
```

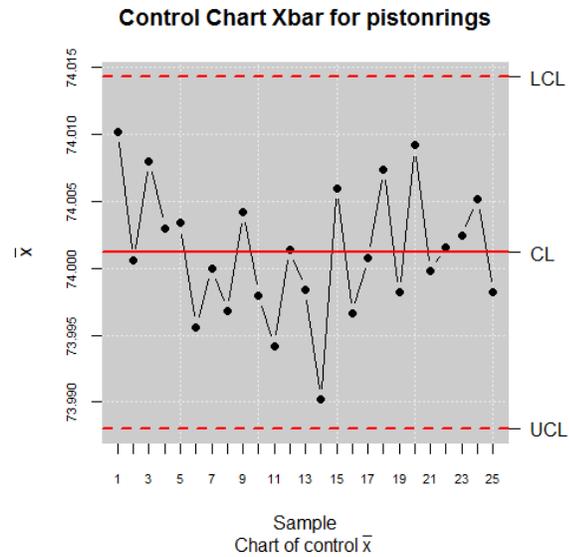


Figura 3.4: Gráfico de Control para las medias de la característica de calidad diámetro (las primeras 25 muestras).

```
'std.dev' is not used when limits is given
> summary(res.qcs)
xbar chart for Control Chart Xbar for pistonrings
```

Summary of group statistics:

```
xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.02
```

```
Group sample size: 5
Number of groups: 40
Center of group statistics: 74.0036
Standard deviation: 0.01007094
```

Control limits:

```
LCL    UCL
```

```
73.98805 74.0143
```

```
Beyond limits of control:
```

```
[1] 74.0166 74.0196 74.0234
```

```
Violationg runs:
```

```
[1] 74.0128
```

```
> plot(res.qcs)
```

Se tiene 4 puntos que determinan el proceso fuera de control (ver Figura 3.5)

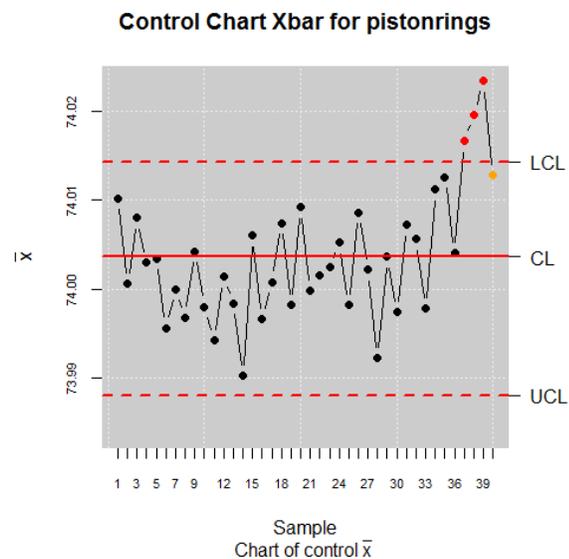


Figura 3.5: Gráfico de Control para las medias de la característica de calidad diámetro (Fase II).

Ejemplo 3: Finalmente se presenta un ejemplo de como realizar el gráfico de control de la variabilidad a través de los rangos (Ver Figura 3.6). Para este ejemplo se utiliza la función `qcs.R`

```
> library(qcr)
> data(pistonrings)
> pistonrings <- droplevels(pistonrings[1:125, ])
> pistonrings.qcd <- qcd(pistonrings)
> res.qcs <- qcs.R(pistonrings.qcd)
```

```
> class(res.qcs)
[1] "qcs.R" "qcs"
> summary(res.qcs)
R chart for pistonrings

Summary of group statistics:
      R
Min.   :0.00800
1st Qu.:0.01800
Median :0.02100
Mean   :0.02276
3rd Qu.:0.02600
Max.   :0.03900

Group sample size: 5
Number of groups: 25
Center of group statistics: 0.02276
Standard deviation: 0.009785039

Control limits:
LCL      UCL
0 0.04812533

Number beyond limits: 0

Number violationg runs: 0
> plot(res.qcs, title = "Control Chart R for pistonrings")
```

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

En este caso la idea es utilizar las desviaciones típicas de cada submuestra en lugar de los rangos para la estimación de la varianza del proceso, por lo que las estimaciones necesarias para la construcción del gráfico de medias consiste en calcular la media y la desviación típica de cada muestra, \bar{x}_i y s_i , respectivamente ($i = 1, \dots, k$). Seguidamente se calcular la media y la desviación típica global. La media global no

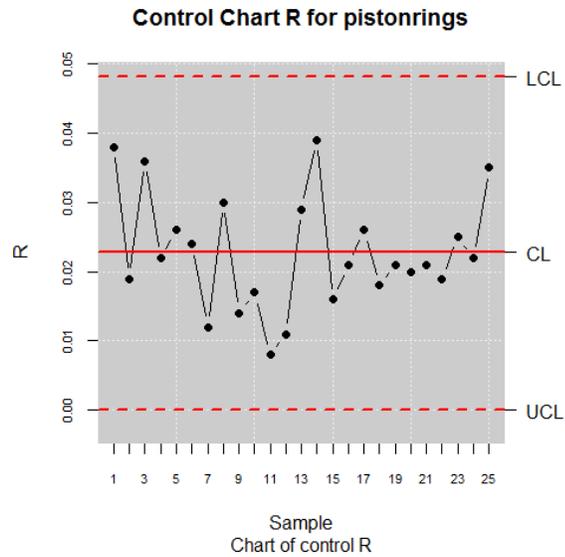


Figura 3.6: Gráfico de Control para los rangos de las 25 muestras de la característica de calidad diámetro.

es más que la media de las medias muestrales, que puede calcularse promediando las medias de cada submuestra o bien promediando el total de los datos (kn).

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \bar{x}_i}{k} = \frac{\sum_{i=1}^k \sum_{j=1}^n x_{ij}}{nk}$$

Esta media resulta ser un estimador insesgado de la media teórica, μ . Como, la desviación típica muestral es un estimador sesgado (aunque asintóticamente insesgado) de la desviación típica teórica. De hecho se tiene $E[s_i] = c_2\sigma$, donde c_2 es una constante, dependiente de n (calculada por el paquete). De esta forma, $\frac{s_i}{c_2}$ es un estimador insesgado de σ y ello motiva el utilizar $\hat{\sigma} = \frac{\bar{s}}{c_2} = \frac{\sum_{i=1}^k s_i}{c_2 k}$ como estimador (centrado) de la desviación típica teórica.

Sin embargo, en el paquete qcr se implementa para la estimación de la varianza las cuasidesviaciones típicas, \hat{s}_i , en lugar de las desviaciones típicas muestrales, en este caso los valores de las constantes cambian. De hecho se tiene $E[\hat{s}_i] = c_4\sigma$, donde c_4 es una constante, dependiente de n . De esta forma, $\frac{\hat{s}_i}{c_4}$ es un estimador insesgado de σ y ello motiva usar:

$$\hat{\sigma} = \frac{\bar{\hat{s}}}{c_4} = \frac{\sum_{i=1}^k \hat{s}_i}{c_4 k}$$

como estimador (centrado) de la desviación típica teórica.

El valor c_4 viene dado por la expresión:

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

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

$$LCL = \bar{\bar{x}} - A_3 \bar{\hat{s}}$$

$$CL = \bar{\bar{x}}$$

$$UCL = \bar{\bar{x}} + A_3 \bar{\hat{s}}$$

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

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

$$LCL = B_3 \bar{\hat{s}}$$

$$LCL = \bar{\hat{s}}$$

$$LCL = B_4 \bar{\hat{s}}$$

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

Ejemplo 4: Se utiliza el mismo conjunto de datos piston rings para realizar un gráfico de medias de los diámetros pero se utiliza en la función `qcs.xbar` la opción `std.dev="UWAVE-SD"` que indica que la estimación de la varianza es por medio de las cuasivarianza de cada subgrupo.

```
> library(qcr)
> data(pistonrings)
> str(pistonrings)
'data.frame': 200 obs. of 3 variables:
 $ diameter: num 74 74 74 74 74 ...
 $ sample : int 1 1 1 1 1 2 2 2 2 2 ...
 $ trial : logi TRUE TRUE TRUE TRUE TRUE TRUE ...
> pistonrings.qcd <- qcd(pistonrings)
>
> class(pistonrings.qcd)
[1] "qcd" "data.frame"
```

```
>
> res.qcs <- qcs.xbar(pistonrings.qcd, std.dev="UWAVE-SD")
> summary(res.qcs)
xbar chart for pistonrings

Summary of group statistics:
      xbar
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.02

Group sample size: 5
Number of groups: 40
Center of group statistics: 74.0036
Standard deviation: 0.01003811

Control limits:
      LCL      UCL
73.99014 74.01707

Beyond limits of control:
[1] 74.0196 74.0234

Violationg runs:
[1] 74.0128
> plot(res.qcs, title = "Control Chart Xbar for pistonrings (UWAVE-SD)")
```

Los límites de control para el gráfico (\bar{x}, R) y el (\bar{x}, \hat{s}) varía a partir del cuarto dígito ya que la estimación de la desviación estándar por el método de los rangos es 0,01007094 y por el método de las cuasidesviaciones es 0,01003811. En la Figura 3.7 se puede observar que se detectan 3 puntos que determinan el proceso fuera de control, que es el mismo resultado obtenido al aplicar el gráfico de control (\bar{x}, R) (Figura 3.1)

Ejemplo 5: En este ejemplo se realiza un gráfico de control para la variabilidad

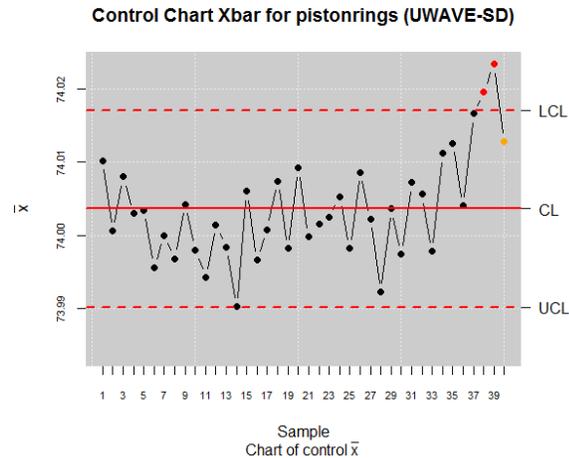


Figura 3.7: Gráfico de Control para las medias utilizando el método de estimación UWAVE-SD”.

utilizando las cuasidesviaciones. Para esto se utiliza la función `qcs.S` del paquete `qcr`.

```
> library(qcr)
> data(pistonrings)
> str(pistonrings)
'data.frame': 200 obs. of 3 variables:
 $ diameter: num 74 74 74 74 74 ...
 $ sample : int 1 1 1 1 1 2 2 2 2 2 ...
 $ trial : logi TRUE TRUE TRUE TRUE TRUE TRUE ...
> pistonrings.qcd <- qcd(pistonrings)
>
> class(pistonrings.qcd)
[1] "qcd" "data.frame"
>
> res.qcs <- qcs.S(pistonrings.qcd)
> class(pistonrings.qcd)
[1] "qcd" "data.frame"
> summary(res.qcs)
S chart for pistonrings
```

Summary of group statistics:

```
S
Min. :0.002864
```

```
1st Qu.:0.007294
Median :0.008703
Mean   :0.009436
3rd Qu.:0.011566
Max.   :0.016547
```

```
Group sample size: 5
Number of groups: 40
Center of group statistics: 0.009435682
Standard deviation: 0.01003811
```

```
Control limits:
LCL          UCL
0 0.01971112
```

```
Number beyond limits: 0
```

```
Number violationg runs: 0
```

Como se puede notar en el gráfico de control de las cuasidesviaciones (Figura 3.8) el proceso con respecto a la variabilidad está bajo control.

Gráficos de medidas individuales

Cuando el tamaño muestral n es 1, se utilizan gráficos de control de medias móviles de distinto orden y para controlar la variabilidad se utiliza gráficos de control de la amplitud móvil entre dos observaciones sucesivas para estimar la variabilidad del proceso.

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

El problema en este tipo de gráfico radica en estimar la varianza, para ello pueden emplearse dos métodos: el de rangos móviles o la agrupación por bloques.

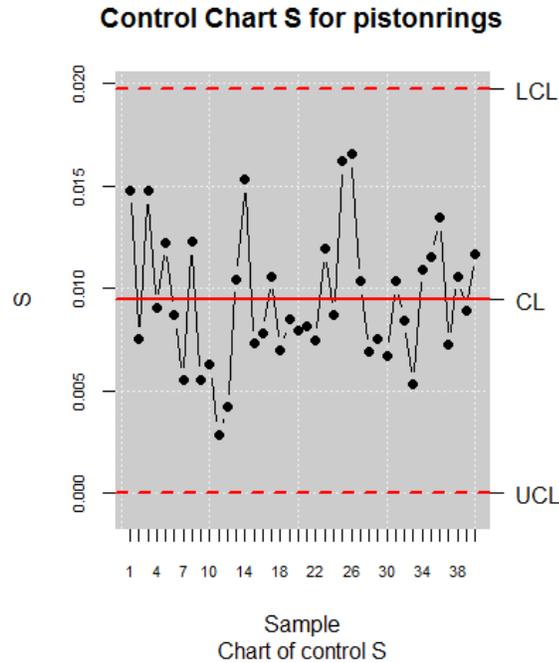


Figura 3.8: Gráfico de Control para las cuasidesviaciones.

El método de rangos móviles consiste en agrupar n medias individuales consecutivas con intención de formar una submuestra que permita estimar los rangos. Una vez creado el primer grupo, se irán agrupando sucesivamente la muestras incorporando la nueva observación y eliminando la primera de cada muestra.

Por otra parte, cuando exista alguna razón que pueda justificar la agrupación de los datos en bloques puede recurrirse a formar los grupos con estos datos. Por ejemplo, en una fábrica de cerveza en la que se tienen dos tanques podría agruparse la producción diaria en 2 bloques formados por la producción diaria en cada tanque.

Para estimar los límites de control en este caso se procede como con el gráfico de la media pero tomando el tamaño de submuestras igual a 1. Sigue utilizándose la misma nomenclatura para el cálculo del estimador de la media, $\bar{\bar{x}}$, en el caso de utilizar gráficos probabilísticos quedarían los límites de control $\bar{\bar{x}} \pm Z_{\frac{\alpha}{2}} \sigma$.

La estimación de σ puede hacerse utilizando la media de las desviaciones típicas \bar{s}_n la media de los rangos \bar{R} , o la media de las cuasidesviaciones \bar{s}_{n-1} :

$$\bar{\bar{x}} \pm Z_{\frac{\alpha}{2}} \frac{\bar{s}_n}{c_{2,n}}$$

$$\bar{\bar{x}} \pm Z_{\frac{\alpha}{2}} \frac{\bar{R}}{d_{2,n}}$$

$$\bar{\bar{x}} \pm Z_{\frac{\alpha}{2}} \frac{\bar{s}_{n-1}}{c_{4,n}}$$

En el supuesto más clásico de usar criterios 3σ suele utilizarse unas constantes E_1 , E_2 y E_3 , resultando las fórmulas para gráficos individuales:

$$\bar{\bar{x}} \pm E_1 \bar{s}_n$$

$$\bar{\bar{x}} \pm \bar{R}$$

$$\bar{\bar{x}} \pm E_3 \bar{s}_{n-1}$$

Ejemplo 6: En este ejemplo se realiza un gráfico de control para observaciones individuales. El ejemplo ha sido tomado del paquete qcc.

```
> library(qcr)
> x <- c(33.75, 33.05, 34, 33.81, 33.46, 34.02, 33.68, 33.27, 33.49, 33.20,
+       33.62, 33.00, 33.54, 33.12, 33.84)
>
> sample <- 1:length(x)
> datos <- data.frame(x,sample)
> datos.qcd <- qcd(datos)
>
> res.qcs <- qcs.one(datos.qcd)
> class(res.qcs)
[1] "qcs.one" "qcs"
> summary(res.qcs)
one chart for datos
```

Summary of group statistics:

```
      one
Min.   :33.00
1st Qu.:33.23
Median :33.54
Mean   :33.52
3rd Qu.:33.78
Max.   :34.02
```

Group sample size: 1

Number of groups: 15
Center of group statistics: 33.52333
Standard deviation: 0.4261651

Control limits:

LCL	UCL
32.24484	34.80183

Number beyond limits: 0

Number violationg runs: 0

```
> plot(res.qcs, title = "Control Chart Xbar.one for pistonrings")
```

Como se puede notar en el gráfico de control de las observaciones individuales (Figura 3.9) determina que el proceso está bajo control.

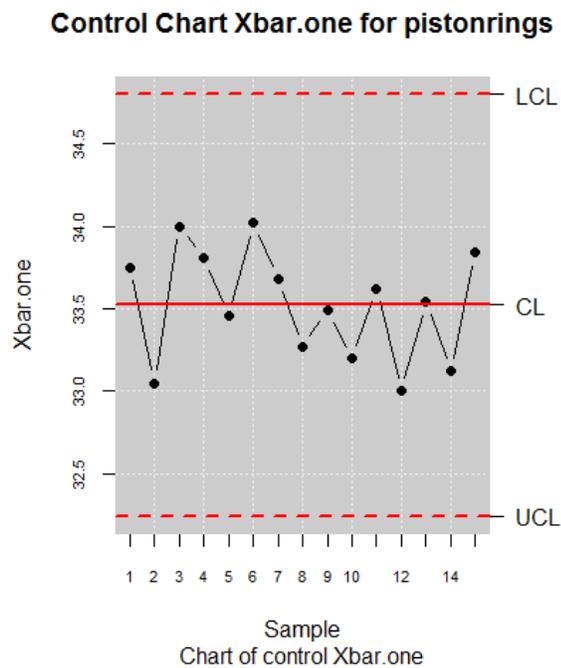


Figura 3.9: Gráfico de Control para observaciones individuales.

3.1.2. El Control de fabricación por atributos

En ocasiones una característica de calidad no puede o no interesa medirse numéricamente y tan sólo se observa si presenta o no determinada propiedad (un producto es defectuoso o no, una pieza encaja o no en otra, un mecanismo funciona o no funciona, etc) que en control de calidad suele emplearse el término conformidad o no conformidad en lugar de éxito o fracaso (defecto). Es este tipo de ejemplos los que se estudiarán en el presente tema bajo el contexto del denominado control por atributos”.

La proporción de defectos o disconformidades es el total de artículos de una determinada población que no cumplen las especificaciones de calidad, siendo esta proporción estimada mediante la toma de muestras aleatorias de la población, mediante la proporción o fracción de disconformidades muestral, que no es más que el cociente entre los artículos disconformes de la muestra y el número total de artículos analizados o tamaño de la muestra elegida.

En esta sección se presentarán los conceptos básicos del control por atributos, analizando los gráficos más habituales como el de fracción de disconformes o gráfico p , el llamado gráfico c que analiza el número de disconformidades y el gráfico u , que se aplica en los estudios en donde interesa estimar el número de defectos o disconformes por unidad.

Para los diferentes ejemplos presentados en esta sección además de la base piston rings se utilizan 3 bases de datos del paquete qcc. A continuación se describe cada una de ellas:

- Datos Jugo de naranja (orangejuice): Es un data frame con 54 Observaciones y 4 variables. Donde se mide el jugo de naranja concentrado congelado que se envasa en latas de cartón de 6 oz Estas latas se forman en una máquina de hilatura. Una lata es inspeccionado para determinar si, cuando se llena, el líquido puede derramarse ya sea en la costura lateral o alrededor de la articulación de la parte inferior. Si esto ocurre, una lata se considera no conforme. Los datos se recogieron como 30 muestras de 50 latas cada una a intervalos de media hora durante un período de tres turnos en el que la máquina estaba en funcionamiento continuo. De la muestra 15 se utiliza en un nuevo batch de stock. Ejemplo 23 se obtuvo cuando un operador sin experiencia fue asignado temporalmente a la máquina. Después de las primeras 30 muestras, se realizó un ajuste de la máquina. Las variables del data frame son:

sample (muestra): Identificación de la muestra

D: número de unidades defectuosas

size (tamaño): tamaño de la muestra

trial (ensayo): muestras de prueba (verdadero / falso)

- Datos Placas impresas (Circuit boards data): Nmero de no conformidades observadas en 26 muestras sucesivas de 100 placas de circuitos impresos. Muestra 6 y 20 están fuera de los límites de control. Muestra 6 fue examinado por un nuevo inspector que no había reconocido varios tipos de no conformidades. Además, el número inusualmente grande de las no conformidades en la muestra 20 se debió a un problema de control de la temperatura en la máquina de soldadura, que fue reparado subsecuentemente. Los últimos 20 muestras son otras muestras recogidas en las unidades de inspección (cada una formada por 100 paneles). Las variables del data frame son:

x : número de unidades defectuosas en 100 placas de circuito impreso (unidad de control)

size (tamaño): tamaño de la muestra

trial (ensayo) indicador muestra ensayo (verdadero / falso)

- Datos del fabricante de ordenadores personales (Personal computer manufacturer data): Un fabricante de ordenadores personales cuenta el número de no conformidades por unidad en la línea de montaje final. Él recoge datos sobre 20 muestras de 5 equipos cada una.

x: número de no conformidades (unidades de control)

size (tamaño): número de equipos inspeccionados

Gráficos de atributos tipo p

El objetivo del gráfico p es controlar la proporción de disconformidades y en la medida de lo posible conseguir que esta proporción disminuya. Así se estudiarán muestras de n elementos, cuyos individuos (artículos) se clasificarán en conformes (correctos) y no conformes (defectuosos).

Si denominamos por p a la proporción de defectuosos del proceso, bajo el supuesto de que esta proporción permanezca estable a lo largo del tiempo y además, dada la hipótesis de independencia, es decir el supuesto de que el hecho de que un artículo

sea defectuoso es independiente de que lo hayan sido o no las anteriores a él, se podrá enfocar el problema de la estimación de la proporción de defectuosos como un proceso de Bernoulli, siendo la distribución del número de defectuosos en una muestra de tamaño n una distribución binomial, que denotaremos como $B(n, p)$.

Hay que recordar que la función de masa de probabilidad para una variable binomial X es:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

Esta variable tiene como media $E(X) = np$, siendo su varianza $Var(X) = npq$.

Para la estimación de la proporción poblacional se elegirá el número de piezas o artículos defectuosos por lote de tamaño n , que denotaremos por \hat{p} , ésta oscilará de una muestra a otra, con media p y desviación típica $\sqrt{\frac{p(1-p)}{n}}$.

En el caso en que el tamaño de las muestras sea suficientemente grande (mayor de 25, aunque es habitual elegir muestras de tamaño 100 o superior) se podría aplicar los resultados que garantizan la aproximación de la distribución binomial por la Normal (teorema central del límite). Utilizando esta última distribución para analizar el proceso.

Los pasos a seguir para la construcción del gráfico de control por atributos o gráfico p serán los siguientes:

- Tomar k muestras (al menos 25) de n elementos ($n > 50$) y contar el número de defectuosos, r_i , en cada muestra i -ésima.
- Estimar la proporción de elementos defectuosos total de las k muestras mediante $\hat{p} = \frac{\sum_{i=1}^k r_i}{nk}$.
- Comprobar que las k muestras son homogéneas respecto de la fracción de elementos defectuosos. Es decir, comprobar si la fracción de defectuosos en cada muestra ($\hat{p}_i = \frac{r_i}{n}$) está dentro de los límites de control

$$\left[\max \left\{ \hat{p} - 3\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, 0 \right\}, \hat{p} + 3\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$$

De esta forma es inmediato reconocer las muestras que no se hallan bajo control. Eliminado del estudio las muestras que están fuera de control y repitiendo los pasos anteriores el número necesario de veces para llegar a un grupo homogéneo puede estimarse fiablemente la capacidad.

- Una vez que todas las muestras se encuentren dentro de los límites, ejecutando los pasos anteriores las veces necesarias, se elegirán esos valores finales que contengan a todas las proporciones muestrales, como límites de control.

Obsérvese que en el caso de que la muestras no sean del mismo tamaño, los límites de los gráficos de control habría que corregirlos por esos tamaños, lo que lleva a que las líneas de los límites de control de los gráficos presenten saltos relacionados con el tamaño de la muestra. Así si llamamos n_i al tamaño de la i -ésima muestra su correspondiente estimación sería:

$$\hat{p}_i = \frac{r_i}{n_i}$$

Ejemplo 7: Se realiza un gráfico de control p para los datos de orangejuice.

```
> library(qcr)
> data(orangejuice)
> str(orangejuice)
'data.frame': 54 obs. of 4 variables:
 $ sample: int  1 2 3 4 5 6 7 8 9 10 ...
 $ D      : int  12 15 8 10 4 7 16 9 14 10 ...
 $ size   : int  50 50 50 50 50 50 50 50 50 50 ...
 $ trial  : logi TRUE TRUE TRUE TRUE TRUE TRUE ...
> datos.qcd <- qcd(data = orangejuice, var.index = 2, sample.index = 1,
+ sizes = size, type.data = "atributte")
> res.qcs <- qcs.p(datos.qcd)
> summary(res.qcs)
p chart for orangejuice
```

Summary of group statistics:

```
      p
Min.   :0.0400
1st Qu.:0.1000
Median :0.1400
Mean   :0.1778
3rd Qu.:0.2400
Max.   :0.4800
```

```

Group sample size: 50
Number of groups: 54
Center of group statistics: 0.1777778
Standard deviation: 0.3823256

```

```
Control limits:
```

```

          LCL          UCL
0.01557078 0.3399848

```

```
Beyond limits of control:
```

```
[1] 0.34 0.44 0.40 0.36 0.48
```

```
Violationg runs:
```

```
[1] 0.18 0.24 0.12 0.04 0.08 0.06 0.12 0.10 0.08 0.16 0.10 0.12 0.14
```

```
[14] 0.10 0.12 0.06 0.10
```

```
> plot(res.qcs)
```

Como se puede notar en el gráfico de control p el proceso está fuera de control (ver Figura 3.10).

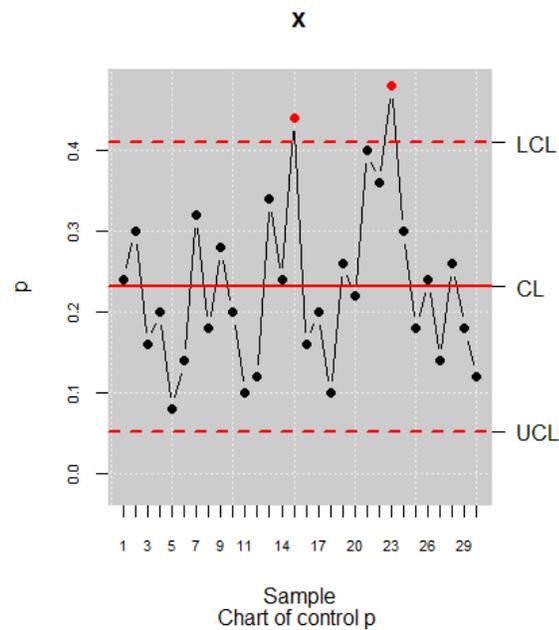


Figura 3.10: Gráfico de Control p .

Gráficos de atributos tipo np

Para la construcción de este tipo de gráficos se estimarán las proporciones de la misma forma que en los gráficos p pero se construirán los gráficos basados en la media del proceso del siguiente modo:

$$LSC_{np} = n\hat{p} + 3\sqrt{n\hat{p}(1-\hat{p})}$$

$$LC_{np} = n\hat{p}$$

$$LIC_{np} = n\hat{p} - 3\sqrt{n\hat{p}(1-\hat{p})}$$

La interpretación del gráfico de control de proporciones (gráfico p) o la del gráfico de número de defectos (gráfico np) es totalmente análoga a la de los gráficos de medias o de dispersión ya tratados en el tema anterior.

Ejemplo 8: Se realiza un gráfico de control np para los datos de orangejuice.

```
> library(qcr)
> data(orangejuice)
> attach(orangejuice)
The following object is masked from circuit (position 3):

  size, trial
The following object is masked from circuit (position 4):

  size, trial
The following object is masked from orangejuice (position 5):

  D, sample, size, trial
The following object is masked from orangejuice (position 6):

  D, sample, size, trial
The following object is masked from orangejuice (position 7):

  D, sample, size, trial
> datos.qcd <- qcd(data = orangejuice, var.index = 2, sample.index = 1,
+ sizes = size, type.data = "atributte")
> res.qcs <- qcs.np(datos.qcd)
> summary(res.qcs)
```

```
np chart for orangejuice
```

```
Summary of group statistics:
```

```
      np
Min.   : 2.000
1st Qu.: 5.000
Median : 7.000
Mean   : 8.889
3rd Qu.:12.000
Max.   :24.000
```

```
Group sample size: 50
```

```
Number of groups: 54
```

```
Center of group statistics: 8.888889
```

```
Standard deviation: 2.70345
```

```
Control limits:
```

```
      LCL      UCL
0.7785388 16.99924
```

```
Beyond limits of control:
```

```
[1] 17 22 20 18 24
```

```
Violating runs:
```

```
[1] 9 12 6 2 4 3 6 5 4 8 5 6 7 5 6 3 5
```

```
> plot(res.qcs)
```

Como se puede notar en el gráfico de control np el proceso está fuera de control (ver Figura 3.11).

Gráficos de atributos tipo c

En este contexto del control de fabricación por número de defectos, se dirá que el proceso está bajo control cuando el proceso es estable (es decir, fabrica un número de defectos por unidad de medida, c , constante a lo largo del tiempo) y los defectos aparecen independientemente los unos de los otros.

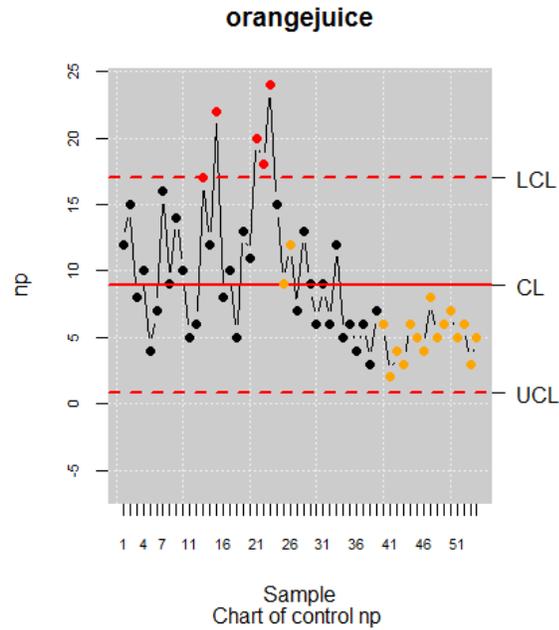


Figura 3.11: Gráfico de Control np.

Suponiendo que el número de defectos por unidad de medida sigue una distribución de Poisson de parámetro c , su función de masa de probabilidad vendrá dada por:

$$P(X = r) = \frac{e^{-c} c^r}{r!}$$

Para estimar c y determinar los gráficos de control, se realizará el siguiente proceso iterativo:

- Se eligen k muestras de tamaño n y se determina el número de defectos de cada una, c_1, c_2, \dots, c_k .
- Se estima c mediante el número medio de defectos por unidad de medida en el total de las k muestras:

$$\hat{c} = \frac{\sum_{i=1}^k c_i}{k}$$

- Se construye un gráfico de control (gráfico c) del número de defectos observados en cada muestra, tomando como línea central \hat{c} y líneas de control $\max\{\hat{c} - 3\sqrt{\hat{c}}, 0\}$ y $\hat{c} + 3\sqrt{\hat{c}}$.

$$LSC_c = \hat{c} + 3\sqrt{\hat{c}}$$

$$LC_c = \hat{c}$$

$$LIC_c = \max \{ \hat{c} - 3\sqrt{\hat{c}}, 0 \}$$

- Se repiten los pasos anteriores el número de veces necesarias, eliminando las muestras que hayan salido de los límites de control.

Ejemplo 9: Se realiza un gráfico de control c para los datos de circuit.

```
> library(qcr)
> data(circuit)
> str(circuit)
'data.frame': 46 obs. of 3 variables:
 $ x      : int  21 24 16 12 15 5 28 20 31 25 ...
 $ size   : int  100 100 100 100 100 100 100 100 100 100 ...
 $ trial: logi  TRUE TRUE TRUE TRUE TRUE TRUE TRUE ...
> datos <- circuit
> datos$sample <- 1:length(datos$x)
> str(datos)
'data.frame': 46 obs. of 4 variables:
 $ x      : int  21 24 16 12 15 5 28 20 31 25 ...
 $ size   : int  100 100 100 100 100 100 100 100 100 100 ...
 $ trial  : logi  TRUE TRUE TRUE TRUE TRUE TRUE TRUE ...
 $ sample: int   1  2  3  4  5  6  7  8  9 10 ...
> sizes <- datos[,2]
> datos.trial<-datos[trial,]
> sizes.trial<-datos[trial,2]
> datos.qcd <- qcd(data = datos, var.index = 1,sample.index = 4,
+ sizes = size, type.data = "atributte")
> res.qcs <- qcs.c(datos.qcd)
> summary(res.qcs)
c chart for datos
```

Summary of group statistics:

```
      c
Min.   : 5.00
1st Qu.:16.00
Median :19.00
```

```

Mean      :19.17
3rd Qu.  :22.00
Max.     :39.00

```

```

Group sample size: 100
Number of groups: 46
Center of group statistics: 19.17391
Standard deviation: 4.378803

```

```

Control limits:
      LCL      UCL
6.037505 32.31032

```

```

Beyond limits of control:
[1] 39 5

```

```

Violations runs:
[1] 12 15
> plot(res.qcs)

```

Como se puede notar en el gráfico de control c el proceso está fuera de control (ver Figura 3.12).

Gráficos de atributos tipo u

En ocasiones puede ser interesante analizar el número medio de disconformidades por unidad de inspección, sin necesidad de que estas unidades contengan el mismo tamaño de piezas analizadas. En este caso se pueden elegir k muestras, suponiendo que no sean del mismo tamaño, y se utilizará la estimación siguiente:

$$\hat{u} = \frac{\sum_{i=1}^k c_i}{\sum_{i=1}^k n_i}$$

donde n_i es el tamaño muestral de la i -ésima muestra, dibujando un nuevo gráfico de control, denominado gráfico u , para el número de defectos por unidad de medida, con línea central en \hat{u} y límites de control para la muestra i fijados en $\max\left\{\hat{u} - 3\sqrt{\frac{\hat{u}}{n_i}}, 0\right\}$ y $\hat{u} + 3\sqrt{\frac{\hat{u}}{n_i}}$ y rechazando toda muestra cuyo número de defectos por unidad de medida, $\frac{c_i}{n_i}$, se salga de los citados límites.

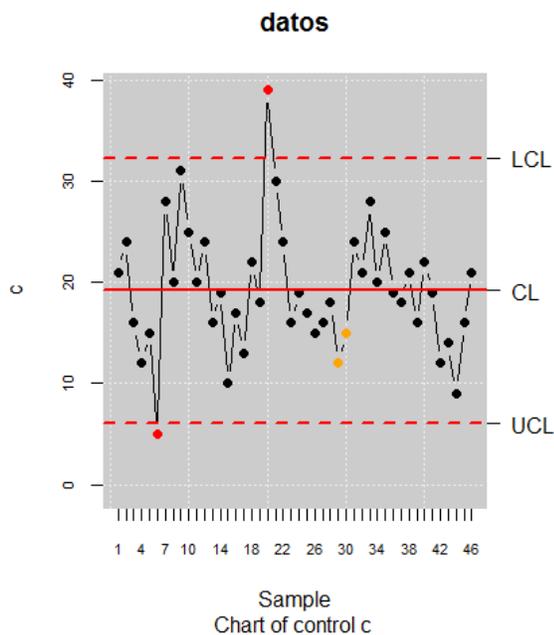


Figura 3.12: Gráfico de Control c.

$$LSC_u = \hat{u} + 3\sqrt{\frac{\hat{u}}{n_i}}$$

$$LC_u = \hat{u}$$

$$LIC_u = \max \left\{ \hat{u} - 3\sqrt{\frac{\hat{u}}{n_i}}, 0 \right\}$$

Ejemplo 10: Se realiza un gráfico de control u para los datos de `pcmanufact`.

```
> library(qcr)
> data(pcmanufact)
> datos <- pcmanufact
> datos$sample <- 1:length(datos$x)
> str(datos)
'data.frame': 20 obs. of 3 variables:
 $ x      : int  10 12 8 14 10 16 11 7 10 15 ...
 $ size   : int  5 5 5 5 5 5 5 5 5 5 ...
 $ sample: int  1 2 3 4 5 6 7 8 9 10 ...
> sizes <- datos[,2]
>
```

```
> datos.qcd <- qcd(data = datos, var.index = 1, sample.index = 3,  
+                 sizes = sizes, type.data = "atributte")  
>  
> res.qcs <- qcs.u(datos.qcd)  
> summary(res.qcs)  
u chart for datos
```

Summary of group statistics:

```
      u  
Min.   :1.00  
1st Qu.:1.40  
Median :2.00  
Mean   :1.93  
3rd Qu.:2.25  
Max.   :3.20
```

```
Group sample size: 5  
Number of groups: 20  
Center of group statistics: 1.93  
Standard deviation: 3.106445
```

```
Control limits:  
      LCL      UCL  
0.06613305 3.793867
```

```
Number beyond limits: 0
```

```
Number violationg runs: 0  
> plot(res.qcs)
```

Como se puede notar en el gráfico de control u el proceso está bajo control (ver Figura 3.13).

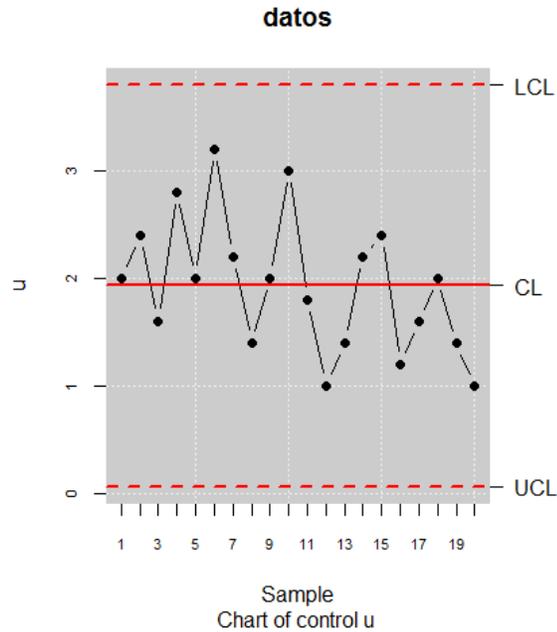


Figura 3.13: Gráfico de Control u.

3.1.3. El Control de fabricación por dependencia

Gráficos CUSUM

El objetivo de este tipo de gráfico de control es controlar la evolución de una variable $X \in N(\mu, \sigma)$, donde σ es conocida o se posee una estimación.

El CUSUM algorítmico (o tabular) calcula las desviaciones de cada valor respecto del valor objetivo o nominal μ . Se distinguirá entre desviaciones positivas y negativas, puesto que en muchas situaciones ambas desviaciones no tienen la misma repercusión.

Es necesario definir un valor K a partir del cual se evidencia que la desviación acumulada es significativa. Este valor determinará la sensibilidad del gráfico de control. Si la suma acumulada hasta la observación i -ésima es menor que cierto umbral K , se considera que la desviación acumulada es cero.

Uniendo esta idea de desviación significativa con las de separación de desviaciones positivas y negativas se tendrán dos tipos de sumas significativas acumuladas: C_i^+ para las desviaciones positivas y C_i^- para las negativas. Estas sumas acumuladas se definen de la siguiente forma:

- C_i^+ = acumulación de desviaciones positivas significativas:

$$C_i^+ = \max \{0, C_i^+ + (X_i - \mu) - K\}$$

De esta forma, si el proceso está bajo control, la variable X tomara valores cercanos al nominal y el termino C_i^+ tendera a tomar valores nulos.

- C_i^- = acumulación de desviaciones negativas significativas:

$$C_i^+ = \min \{0, C_i^- - (X_i - \mu) - K\}$$

El valor de K se suele elegir en función de la desviación que se quiera detectar. Supongamos que el valor nominal de la variable de interés es μ .

Los valores de las sumas acumuladas se han de comparar con el valor máximo admisible. Este valor, denominado valor de decisión H suele ser $H = h\sigma$, donde se recomienda que $h = 5$ y por tanto $H = 5\sigma$ (algunos autores también recomiendan valores de $h = 4$). Los valores $\pm H$ son los límites de control de este gráfico.

Ejemplo 11: Se realiza un gráfico de control *cusum* para los datos de pistonrings.

```
> library(qcr)
> data(pistonrings)
> res.qcd <- qcd(pistonrings, type.data = "dependence")
> res.qcs <- qcs.cusum(res.qcd, type = "cusum")
> summary(res.qcs)
cusum chart for pistonrings
```

Summary of group statistics:

```
cusum
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.01
Max.   :74.02
```

```
Group sample size: 5
Number of groups: 40
Center of group statistics: 74.0036
Standard deviation: 0.01007094
```

Control limits:

```
[1] -5 5
```

Beyond limits of control:

```
[1] 73.9902 74.0060 73.9966 74.0008 74.0074 73.9982 74.0092 73.9998 74.0016
[10] 74.0024 73.9982 73.9922 73.9974
```

Violationg runs:

```
[1] 74.0196 74.0234 74.0128
> plot(res.qcs)
```

Como se puede notar en el gráfico de control *cusum* el proceso está fuera de control(ver Figura 3.14).

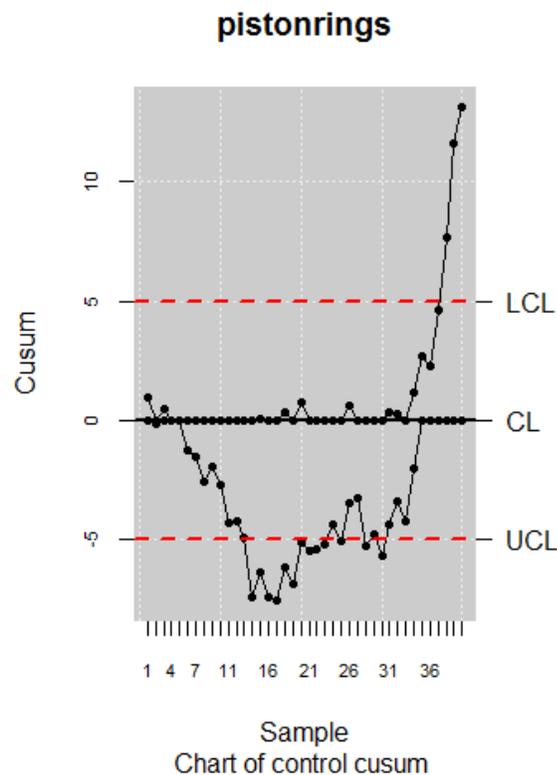


Figura 3.14: Gráfico de Control CUSUM.

Gráficos EWMA

La salida de un proceso bajo supervisión puede analizarse también mediante la filosofía de las series temporales, pues en el fondo cuando se hacen las correspondien-

tes gráficas de las submuestras en distintos períodos de tiempo, estas constituyen una serie temporal. En este sentido es donde se podría englobar los gráficos presentados en esta sección.

El gráfico de medias móviles con ponderación exponencial o gráfico EWMA (Exponentially Weighted Moving Average), fue introducido primeramente por Roberts (1959) y luego por Wortham y Ringer (1971), que lo propusieron para aplicarlo en procesos industriales, en sistemas de control en finanzas y en dirección, para las que los subgrupos formados por varios elementos no son aplicables.

Para este tipo de gráficos se toman normalmente datos individuales. Las observaciones individuales pueden ser medias (cuando las observaciones individuales de las que provienen las medias no están disponibles), lecturas individuales, cocientes, proporciones, o medidas similares.

El estadístico dibujado es la media ponderada de la observación actual y todas las observaciones precedentes, donde la media previa recibe más peso, es decir, $Z_t = \lambda x_t + (1 - \lambda) Z_{t-1}$, $0 < \lambda < 1$.

Donde: $Z_0 = \mu$, Z_t y Z_{t-1} son las medias móviles ponderadas exponencialmente en el tiempo actual t y $t - 1$ respectivamente, x_t es la observación presente y λ es el factor de ponderación para la observación actual.

Si se cumple la hipótesis de idenpendencia y la estabilidad sobre los x_t , se tiene que $E(Z_t) = \mu$ y $Var(Z_t) = \sigma^2 \frac{\lambda(1-\lambda)^{2t}}{2-\lambda}$

Cuando t aumenta, el último termino entre paréntesis en la parte derecha converge rápidamente hacia uno, y la correspondiente expresión para la varianza se transforma en $Var(Z_t) = \sigma^2 \left[\frac{\lambda}{2-\lambda} \right]$.

Eligiendo $\lambda = \frac{2}{t+1}$, la expresión aproximada de la varianza se transforma en $Var(Z_t) \approx \frac{\sigma^2}{t}$ (varianza de medias de muestras de t elementos).

Bajo estas condiciones, los límites de control se transforman en $\hat{\mu} \pm 3\sqrt{\frac{\hat{\sigma}^2}{t}}$. Para otros valores de λ , los límites de control son:

$$LSC = \hat{\mu} + 3\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}}$$

$$LSC = \hat{\mu} - 3\hat{\sigma}\sqrt{\frac{\lambda}{2-\lambda}}$$

Para las primeras observaciones, se debe utilizar la primera ecuación para la varianza. Si no se tiene una buena estimación para σ , se debe utilizar un gráfico de recorridos y estimar por $\frac{\bar{R}}{d_2\sqrt{n}}$. En el caso de valores individuales, el recorrido móvil medio se puede utilizar en la forma en que se hizo para los gráficos de control de

valores individuales.

El gráfico EWMA puede expresarse de la forma general:

$$Z_{t,k} = \sum_{i=1}^k (w_i \bar{x}_{t-i+1}) + a$$

En donde t es el instante actual, k es el número de medias consideradas en cada punto y que se llama orden del gráfico, w_i son los k pesos considerados y a es una constante que puede ser nula. Con esta expresión general pueden obtenerse distintos gráficos:

- Si $k = 1$, $w_i = 1$ y $a = 0$ obtendríamos el gráfico de Walter A. Shewhart para la media.
- Si $w_i = 1$ y $a = -k\mu_0$ se obtiene el gráfico CUSUM en donde:

$$S_{t,k} = \sum_{i=1}^k (\bar{x}_{t-i+1} - \mu_0) = \sum_{i=1}^k \bar{x}_{t-i+1} - k\mu_0$$

- Si $w_i = \frac{1}{k}$ y $a = 0$ se obtiene el denominado gráfico de Medias Móviles

$$M_{t,k} = \frac{1}{k} \sum_{i=1}^k \bar{x}_{t-i+1}$$

Si $w_i = \lambda(1 - \lambda)^{i-1}$ y $a = (1 - \lambda)^t \mu_0$ es el gráfico EWMA donde el estadístico está expresado como:

$$Z_t = \sum_{i=1}^k \lambda(1 - \lambda)^{i-1} \bar{x}_{t-i+1} + (1 - \lambda)^t \mu_0$$

Tanto el gráfico CUSUM y el gráfico EWMA son más efectivo que el gráfico de medias para detectar pequeños cambios en la media; sin embargo, ambos gráficos funcionan peor que el gráfico de medias para grandes cambios.

Ejemplo 12: Se realiza un gráfico de control *ewma* para los datos de pistonrings.

```
> library(qcr)
> data(pistonrings)
> x <- droplevels(pistonrings[1:125,])
> res.qcd <- qcd(x, type.data= "dependence")
```

```
> res.qcs <- qcs.ewma(res.qcd)
> summary(res.qcs)
ewma chart for x
```

Summary of group statistics:

```
      ewma
Min.   :73.99
1st Qu.:74.00
Median :74.00
Mean   :74.00
3rd Qu.:74.00
Max.   :74.01
```

```
Group sample size: 5
Number of groups: 25
Center of group statistics: 74.00118
Standard deviation: 0.009785039
```

Control limits:

```
      LCL      UCL
1 73.99855 74.00380
2 73.99781 74.00454
3 73.99742 74.00493
4 73.99718 74.00517
5 73.99704 74.00531
6 73.99695 74.00540
7 73.99690 74.00545
8 73.99686 74.00549
9 73.99684 74.00551
10 73.99683 74.00553
11 73.99682 74.00554
12 73.99681 74.00554
13 73.99681 74.00555
14 73.99680 74.00555
15 73.99680 74.00555
16 73.99680 74.00555
```

```

17 73.99680 74.00555
18 73.99680 74.00555
19 73.99680 74.00555
20 73.99680 74.00555
21 73.99680 74.00555
22 73.99680 74.00555
23 73.99680 74.00555
24 73.99680 74.00555
25 73.99680 74.00555
> plot(res.qcs, title = "Control Chart EWMA")

```

Como se puede notar en el gráfico de control *ewma* el proceso está fuera de control (ver Figura 3.15).

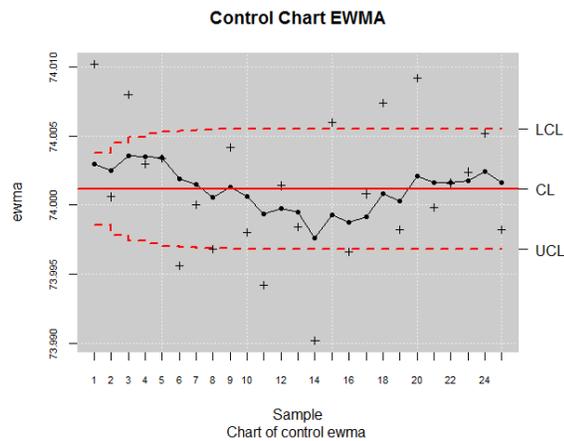


Figura 3.15: Gráfico de Control EWMA.

3.2. Gráficos de control multivariante

3.2.1. Control de proceso multivariantes basados en el supuesto de normalidad

En esta sección se utilizan los conjuntos de datos del paquete MSQC que deben ser de tipo array. Este tipo de dato es una generalización del tipo de dato data frame o matriz que se maneja en R. En el caso de que no sea un array el paquete qcr lo transforma de manera automática a array. A continuación se presentan los conjuntos de datos que se utilizan para los ejemplos.

- Pasador de una clavija (Dowel pin): Diámetro y longitud de un proceso de fabricación de un pasador de clavija.

diamter (diámetro): un vector numérico

longitud (length): un vector numérico

- Target archery: Conjunto de datos de tiro con arco en la ronda de clasificación. En esta etapa el arquero dispara 72 flechas en 12 fines de seis flechas. La información que se da en las coordenadas x e y. Es un array de dimensiones 24 x 2 x 3.

x-coordinate: coordenada x

y-coordinate: coordenada y

La T^2 Hotelling

Como ya se había indicado Hotelling en 1947 es el que impulsa y realiza el primer trabajo en control de calidad multivariado cuya aplicación fue a datos sobre precisión de bombas durante la segunda guerra mundial. El contexto en el que se desarrollan los gráficos de control de calidad multivariado es el siguiente.

Consideremos un proceso de control en el que se observan p variables que se distribuyen según una Normal p variante con un vector de medias $\mu_0 = (\mu_{01}, \mu_{02}, \dots, \mu_{0p})$ y una matriz Σ_0 de varianzas-covarianzas de orden p dentada por $N_p(\mu_0, \Sigma_0)$. Se desea contrastar las hipótesis: $H_0 : \mu_i = \mu_0$ vs. $H_1 : \mu_i \neq \mu_0$ para determinar si el proceso está bajo control. Para esto se diseñan criterios para determinar si el proceso esta fuera de control que utilizan el estadístico T^2 de Hotelling que dependiendo del estado de proceso sigue una determinada distribución.

- Cuando el proceso está bajo control, ($\mu_i = \mu_0$), existe una probabilidad α de que el estadístico T_i^2 exceda al valor crítico $\mathcal{X}_{p,\alpha}^2$. Por ello se toma como indicación o señal de fuera de control un valor del estadístico T_i^2 que supera al valor $\mathcal{X}_{p,\alpha}^2$.

El estadístico T^2 Hotelling es:

$$T_i^2 = n (\bar{X}_i - \mu_0)^t \Sigma_0^{-1} (\bar{X}_i - \mu_0)$$

donde $\bar{X}_i = (\bar{X}_{i1}, \dots, \bar{X}_{ip})$ es el vector de medias para las p -características, en la i -ésima muestra. Cuando el tamaño del subgrupo $n = 1$, se obtendrán vectores de observaciones en lugar de vectores de medias.

En el contexto geométrico, el estadístico T^2 es una proporción de la distancia (Mahalanobis) al cuadrado entre una observación multivariada y el vector de valores objetivo (vector de media), donde puntos equidistantes forman elipsoides alrededor de dicho vector. A mayor valor de T^2 , mayor es la distancia entre la observación y el valor objetivo.

Si μ_0 y Σ_0 son conocidos y el proceso está bajo control, el estadístico T_i^2 se distribuye como una chi-cuadrado central con p grados de libertad ($T_i^2 \approx \chi_p^2$).

- Cuando el proceso no está bajo control, el estadístico T_i^2 se distribuye como una variable chi-cuadrado no central con p grados de libertad y con parámetro de no centralidad: $\lambda = n(\mu_1 - \mu_0)^t \Sigma_0^{-1} (\mu_1 - \mu_0)$, con $\mu_1 \neq \mu_0$; donde μ_1 es el vector de medias de las p características, cuando existe un cambio d en al menos una de las medias.

En la práctica, el valor del estadístico T^2 se estimará mediante los correspondientes valores muestrales, dando como resultado la siguiente expresión:

$$T^2 = n (\bar{X} - \bar{\bar{X}})^t S^{-1} (\bar{X} - \bar{\bar{X}})$$

Con distribución bajo la hipótesis nula:

$$T^2 \approx \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

Es decir, el estadístico T^2 sigue aproximadamente en este caso una F de Snedecor con p y $(mn - m - p + 1)$ grados de libertad, para un nivel de significación del α %.

Dado que tanto la distribución χ^2 y la F son asimétricas con sesgo a la derecha, se va a considerar únicamente la expresión para el límite de control superior, que viene dada por:

$$LCS = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$

Para el límite inferior se considera $LCI = 0$

Además, para procesos en los que el m es suficientemente grande, $m > 100$, se puede aproximar este límite por $LCS = X_{\alpha,p}^2$.

Alt (1985) recomienda usar estos límites en una primera fase, lo que denomina análisis retrospectivo, para una vez estimado el límite de control por este medio, calcular un segundo límite para la segunda fase, la de monitorización del proceso, para

ello propone un cambio consistente en multiplicar por la expresión $\frac{m+1}{m-1}$ obteniendo así una nueva expresión para el límite de control superior LCS.

En el caso de observaciones individuales, es decir $n = 1$, suponiendo que se tienen un total de m muestras, para evaluar p variables de control, el correspondiente gráfico de control se construirá mediante:

$$LCS = \frac{p(m+1)(n-1)}{m^2 - mp} F_{\alpha, p, m-p}$$

$$LCI = 0$$

Los dos problemas principales que surgen al aplicar un gráfico T^2 son:

- La escala que se utiliza para graficar los valores, no está relacionada con la escala de ninguna de las variables involucradas.
- Cuando el estadístico T^2 excede el límite superior de control (L.S.C), el usuario no posee información relevante sobre qué pudo haberla causado. Se requiere información adicional para determinar que variables son responsables de ese valor fuera de límite de control. Esta información es esencial en cualquier intento de detectar las causas de las desviaciones y posiblemente corregir el proceso. Por eso aconsejan llevar en forma paralela, gráficos de control univariados para cada característica, utilizar métodos como los propuestos por Mason, Tracy y Young (1995) o Kourty y Mc Gregor (1996).

Ejemplo 13: Se realiza un gráfico de control T^2 para los datos de dowel1.

```
> library(qcr)
> data(dowel1)
> str(dowel1)
'data.frame': 40 obs. of 2 variables:
 $ diameter: num  0.492 0.501 0.491 0.492 0.505 0.5 0.497 0.509 0.49 0.499 ...
 $ length  : num  0.988 1.011 1.008 0.97 1.003 ...
> data.mqcd <- mqcd(dowel1)
> res.mqcs <- mqcs.t2(data.mqcd)
> summary(res.mqcs)
t2 chart for dowel1
Summary of group statistics:
      t2
Min.   :0.09137
```

```

1st Qu.:0.60154
Median :1.66096
Mean   :1.95000
3rd Qu.:2.64402
Max.   :5.34020

```

```

Number of quality characteristics: 2
Number of samples or observations: 40
Number of observations or sample size: 1

```

Mean Vector:

```
0.500875 1.001825
```

Covariance Matrix:

```

           diameter      length
[1,] 4.908654e-05 8.584936e-05
[2,] 8.584936e-05 4.199429e-04

```

Control limits:

```

      lcl      ucl
0.00000 12.44888

```

Number beyond limits: 0

```
> plot(res.mqcs, title = " Hotelling Control Chart for dowel1")
```

Como se puede notar en el gráfico de control T^2 el proceso está bajo control(ver Figura 3.16).

Ejemplo 14: Se realiza un gráfico de control T^2 para los datos de archery1.

```

> library(qcr)
> data(archery1)
> str(archery1)
 num [1:24, 1:2, 1:3] 24.14 28.55 3.97 28.57 -3.43 ...
 - attr(*, "dimnames")=List of 3
  ..$ : NULL
  ..$ : chr [1:2] "x-coordinate" "y-coordinate"
  ..$ : NULL
> data.mqcd <- mqcd(archery1)

```

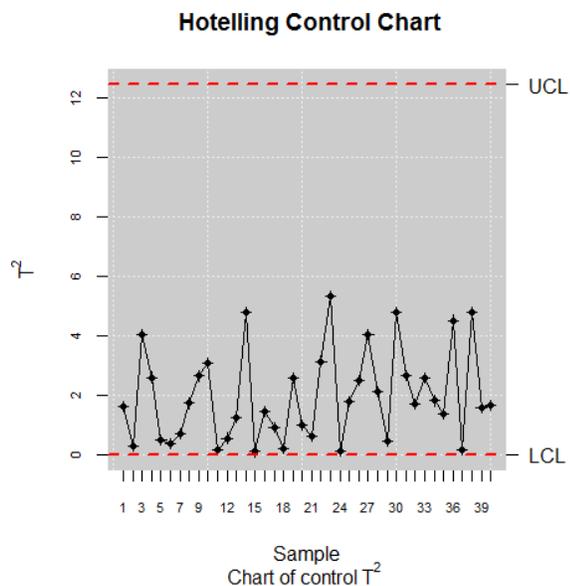


Figura 3.16: Gráfico de Control T^2 para los datos de dowell1.

```
> res.mqcs <- mqcs.t2(data.mqcd)
> summary(res.mqcs)
t2 chart for archery1
```

Summary of group statistics:

```
      t2
Min.   :0.04769
1st Qu.:0.37341
Median :0.92883
Mean   :1.51997
3rd Qu.:2.11386
Max.   :6.16892
```

Number of quality characteristics: 2

Number of samples or observations: 24

Number of observations or sample size: 3

Mean Vector:

```
 6.779028 5.772917
```

Covariance Matrix:

```

      [,1]      [,2]
[1,] 105.25999 48.44271
[2,] 48.44271 149.28805

```

Control limits:

```

      lcl      ucl
0.000000 9.958262

```

Number beyond limits: 0

```
> plot(res.mqcs, title = " Hotelling Control Chart for archery1")
```

Como se puede notar en el gráfico de control T^2 el proceso está bajo control (ver Figura 3.17).

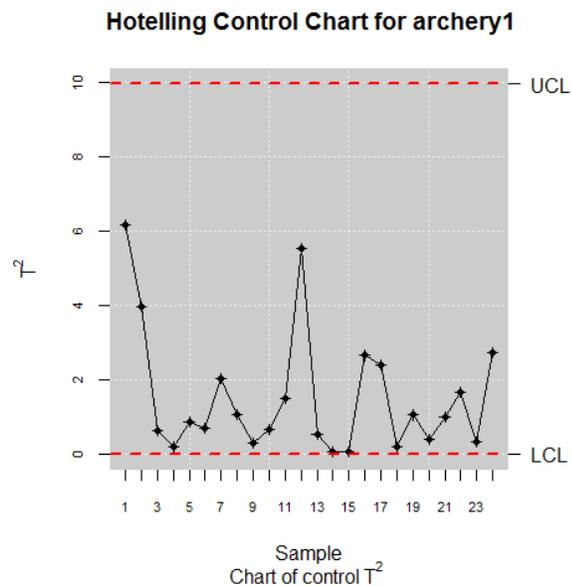


Figura 3.17: Gráfico de Control T^2 para los datos de archery1.

Gráficos MCUSUM

Los gráficos CUSUM pueden extenderse también al caso multivariante, aunque no hay una única forma de hacerlo.

Los primeros en realizar estudios en la materia fueron Woodall y Ncube (1985), usando un esquema basado en múltiples (p) CUSUM univariantes.

Healy (1987) desarrolló un modelo de gráfico CUSUM multivariante mediante:

$$G_i = \max \left\{ \left(G_{i-1} + a^t (x_i - \mu_0) - 0, 5\lambda\mu_1 \right), 0 \right\}$$

Donde:

μ_0 : Vector de medias cuando el proceso está bajo control

Σ_0 : Matriz de varianzas-covarianzas bajo control

μ_1 : Vector de medias cuando el proceso no está bajo control

$\lambda\mu_1$: Raíz cuadrada del parámetro de descentrado

$$a^t = \frac{A}{\lambda\mu_1}, \quad A = (\mu_1 - \mu_0)^t \Sigma_0^{-1}$$

Para interpretar el gráfico se considera que se produce una señal de alarma cuando G_i supera un cierto valor de referencia H , convenientemente elegido y que se encuentra tabulado. Hawkins (1991) desarrolla un modelo para variables ajustadas mediante regresión multivariante, en base a la suposición de que cuando se produce un cambio en la media, se produce de una determinada manera. Concretamente se asume que el cambio se debe a una variación de magnitud d en la media de una de las variables.

Otras propuesta de gráficos MCUSUM son los de Croisier (1988) o los de Pignatiello y Runger (1990). En el paquete qcr se ha implementado el primer procedimiento de Croisier que reduce cada observación multivariada a un escalar y luego construye el estadístico CUSUM con los escalares.

Croisier considera el siguiente estadístico:

$$T_i = \sqrt{T_i^2} = \sqrt{n (\bar{X}_i - \mu_0)^t \Sigma_0^{-1} (\bar{X}_i - \mu_0)}$$

Croisier considera el estadístico T porque de esa manera, se acumulan distancias en lugar de distancias al cuadrado. El estadístico MCUSUM se calcula de la siguiente manera:

donde $S_0 \geq 0$ (en general se considera $S_0 = 0$) y $k > 0$.

El gráfico MCUSUM da una señal de fuera de control cuando el valor de S_i es mayor que un cierto valor h que depende del valor del ARL deseado cuando el proceso está funcionando en el valor objetivo [$S_i > h$].

Para determinar los valores de h y k para $p = 2, 5, 10$ y 20 y para valores de ARL de 200 y de 500 cuando el proceso está bajo control Croisier empleo procesos de Markov. Estas tablas se diseñaron para detectar un cambio $d = 1$ en el vector de medias, donde $d = \lambda$, con λ parámetro de no centralidad. En este caso especial los valores de k óptimos se aproximan a p .

Ejemplo 15: Se realiza un gráfico de control MCUSUM para los datos de dowel1.

```
> library(qcr)
> data(dowell1)
> str(dowell1)
'data.frame': 40 obs. of 2 variables:
 $ diameter: num  0.492 0.501 0.491 0.492 0.505 0.5 0.497 0.509 0.49 0.499 ...
 $ length  : num  0.988 1.011 1.008 0.97 1.003 ...
> data.mqcd <- mqcd(dowell1)
> res.mqcs <- mqcs.mcusum(data.mqcd)
> res.mqcs <- mqcs.mcusum.default(dowell1)
> summary(res.mqcs)
```

Summary of group statistics:

```
      t2
Min.   :0.0000
1st Qu.:0.7749
Median :1.1501
Mean   :1.3818
3rd Qu.:1.9987
Max.   :3.8409
```

```
Number of quality characteristics: 2
Number of samples or observations: 40
Number of observations or sample size: 1
```

Mean Vector:

```
0.500875 1.001825
```

Covariance Matrix:

```
      diameter      length
[1,] 4.908654e-05 8.584936e-05
[2,] 8.584936e-05 4.199429e-04
```

Control limits:

```
lcl ucl
0.0 5.5
```

Number beyond limits: 0

```
> plot(res.mqcs, title = " MCUSUM Control Chart for dowel1")
```

Como se puede notar en el gráfico de control MCUSUM el proceso está bajo control(ver Figura 3.18).

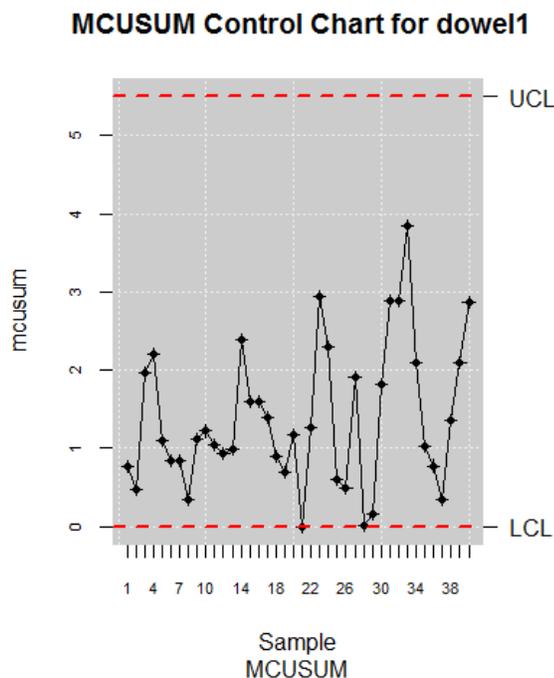


Figura 3.18: Gráfico de Control MCUSUM para los datos de dowel1.

Gráficos MEWMA

El gráfico de medias móviles ponderadas exponencialmente para el caso multivariante (MEWMA) es una extensión a esta nueva situación del correspondiente gráfico univariante. La extensión multivariante, propuesta por Lowry (1992), adopta la forma:

$$Z_i = \Lambda X_i + (I - \Lambda) Z_{i-1}$$

donde: X_i es el vector de medias muestrales y Λ es la matriz diagonal formada por los valores λ para las distintas variables. Los λ_j marcan la profundidad de la memoria para cada variable. A mayor valor de r_j , menor profundidad. I es la matriz identidad y se considera como valor inicial $Z_0 = 0$.

La información que proporcionan los Z_i se recoge en el estadístico:

$$T_i^2 = Z_i^t \Sigma_{Z_i}^{-1} Z_i$$

Donde $\Sigma_{Z_i}^{-1}$ es la inversa de la matriz de varianzas-covarianzas de los Z_i .

La señal de salida de control se produce cuando T_i^2 supera un cierto valor h , ($h > 0$) seleccionado de manera tal de lograr un cierto valor de ARL cuando el proceso está bajo control. Si no existe a priori ninguna razón para ponderar en forma diferente las observaciones pasadas de cada una de las p variables (como generalmente sucede), entonces se considera $\lambda_1 = \dots = \lambda_p$.

La matriz Σ_{Z_i} puede obtenerse a partir de los elementos de la matriz de varianzas-covarianzas correspondiente a las variables analizadas mediante la expresión:

$$\Sigma_{Z_i} = \frac{\lambda}{2 - \lambda} [1 - (1 - \lambda)^{2i}] \Sigma_x$$

donde Σ_x es la matriz de varianzas-covarianzas original.

Cuando $r = 1$, el gráfico MEWMA coincide con el gráfico de control T^2 dado que el valor asintótico de la matriz de varianzas-covarianza de Z_i es:

$$\Sigma_{Z_i} = \frac{\lambda}{2 - \lambda} \Sigma_x$$

En lo que respecta al límite de control (superior) empleado, Runger y Prabhu (1996) sugieren una aproximación mediante cadenas de Markov, que permite estudiar el funcionamiento del gráfico referente al ARL . Proporcionan además una serie de recomendaciones para la selección de los parámetros del gráfico.

En lo que se refiere al parámetro λ , Montgomery (1991) recomienda que su valor está comprendido entre 0.05 y 0.25. Cuanto mayor sea el valor del parámetro, menor importancia se le estará dando a los valores más alejados en el tiempo. En aplicaciones prácticas se elige 0.1 como el valor del parámetro λ .

Ejemplo 16: Se realiza un gráfico de control MEWMA para los datos de dowel1.

```
> library(qcr)
Summary of group statistics:
      t2
Min.   :0.04165
1st Qu.:0.95586
Median :1.69276
Mean   :1.62204
3rd Qu.:1.89606
```

```
Max.      :4.17595

Number of quality characteristics:  2
Number of samples or observations:  40
Number of observations or sample size:  1

Mean Vector:
  0.500875  1.001825
Covariance Matrix:
      diameter      length
[1,] 4.908654e-05 8.584936e-05
[2,] 8.584936e-05 4.199429e-04

Control limits:
  lcl    ucl
0.0000  8.6336

Number beyond limits: 0
> plot(res.mqcs, title = " MEWMA Control Chart for dowel1")
```

Como se puede notar en el gráfico de control MCUSUM el proceso está bajo control(ver Figura 3.19).

3.3. Interfaz gráfica para el control de calidad de procesos industriales

En esta sección se presentan los cuatro módulos (variable, atributo, dependencia y multivariante) desarrollados con el paquete Shiny para el control de calidad de procesos industriales. Cada uno de estos módulos mantienen la misma estructura, cada módulo consiste de dos paneles, un panel de configuración en el cual se puede seleccionar el conjunto de datos a utilizar; el tipo de gráfico de control a implementar; el número de desviaciones estandar (3 sigmas) a utilizar para los límites de control y el título del gráfico. Adicionalmente, se puede especificar el número de observaciones a visualizar y en el caso que se desee descargar el conjunto de datos se tiene el botón Download para esta acción. En la Figura 3.20 se presenta el panel de configuración para el módulo de gráficos de control para variables cuantitativas.

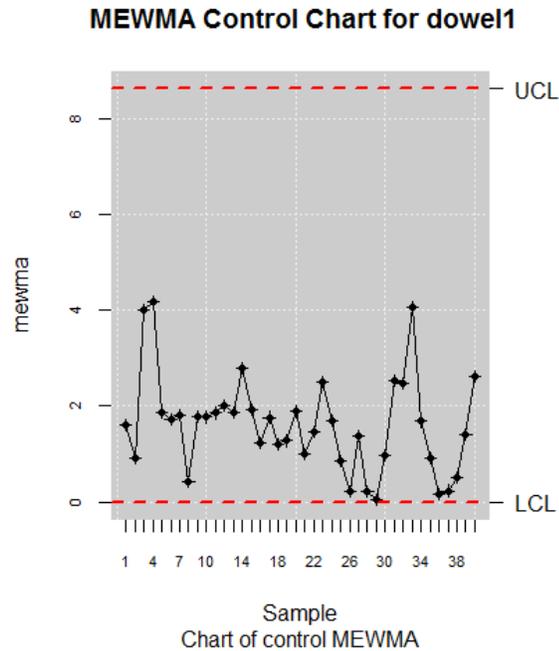


Figura 3.19: Gráfico de Control MEWMA para los datos de dowel1.

Para el modulo de gráficos de control de variables cuantitativas se tiene dos conjunto de datos “pistonrings” y “one”. El primero se utiliza para realizar los gráficos del tipo \bar{x} , R y S; mientras que el segundo conjunto de datos para el gráfico de control de medidas individuales. En la Figura 3.20 se ha seleccionado el conjunto de datos “pistonrings”, y sólo se visualizan las 10 primeras observaciones de estos datos.

El segundo panel es de visualización de los resultados y consiste en tres entornos para observar información del conjunto de datos (Figura 3.22), el gráfico de control (ver Figura 3.23) y la estadística de resumen (ver Figura 3.24). La información presentada en el entorno gráfico se genera de forma dinámica directamente desde R y a partir del paquete qcr. En la Figura 3.23 las líneas punteadas de color rojo son los límites de control y la continua es el valor central del estadístico. Los límites de control están determinados a 3 sigmas, en el caso que se quisiera cambiar los límites de control se puede especificar mediante la opción de número de desviaciones estandar (Figura 3.21).

En la Figura 3.23 y Figura 3.24 se puede apreciar el gráfico de control \bar{x} y el resumen de los estadísticos generados a partir del paquete qcr. Una de las ventajas de utilizar la interfaz gráfica es poder visualizar la información sin tener que interactuar directamente con las funciones creadas en el paquete qcr y además tener la facilidad

Figura 3.20: Panel donde se especifica el tipo de gráfico de control a implementar de poder contar con una salida gráfica y estadísticas de resumen en un ambiente web.

Figura 3.21: Número de desviaciones estandar

Por ejemplo, es muy fácil interactuar entre el gráfico que se presenta en la Figura 3.23 donde se tiene 3 puntos que determinan que el proceso no está bajo control y el resultado de la Figura 3.24 en la que se tiene un resumen del estadístico (que en este caso es la media) y los límites de control. Mediante el gráfico se detecta los puntos y mediante el resumen se tiene los valores de los tres puntos que se observan en la Figura 3.23.

Todos módulos tienen el mismo esquema de funcionamiento la única diferencia entre los módulos radica en el conjunto de datos a seleccionar y el tipo de gráfico de control a implementar. En el Apendice A, se presentán los modulos para el caso de variables cuantitativas, cualitativas, dependientes y el caso multivariante.

A continuación para cada módulo se presentan los tipos de gráficos de control que están implementados a partir del paquete qcr.

- Módulo 1 (cuantitativas): xbar, S, R y one
- Módulo 2 (cuanlitativas): p, np, c y u

View Plot Summary

```

diameter      sample      trial
Min.   :73.97  Min.    : 1.00  Mode :logical
1st Qu.:74.00  1st Qu.:10.75  FALSE:75
Median :74.00  Median :20.50  TRUE :125
Mean   :74.00  Mean   :20.50  NA's :0
3rd Qu.:74.01  3rd Qu.:30.25
Max.   :74.04  Max.   :40.00

```

	diameter	sample	trial
1	74.03	1	TRUE
2	74.00	1	TRUE
3	74.02	1	TRUE
4	73.99	1	TRUE
5	74.01	1	TRUE
6	74.00	2	TRUE
7	73.99	2	TRUE
8	74.00	2	TRUE
9	74.01	2	TRUE
10	74.00	2	TRUE

Figura 3.22: Resumen del conjunto de datos pistonrings

- Módulo 3 (dependientes): cusum y ewma
- Módulo 4 (multivariante paramétrico): t^2 , mcusum y mewma

En cada módulo en el panel de configuración se puede aumentar más parámetros del proceso, es decir se puede dejar que el usuario ingrese una media nominal y una varianza poblacional, así también parámetros de calibrado para la construcción de los límites de control. Por otro lado, dado el carácter modular del paquete y de la interfaz gráfica aumentar sus funcionalidades es muy fácil ya que ningún módulo es dependiente de otro módulo y por tanto no se requiere del conocimiento de como se encuentra desarrollado los otros módulos para crear nuevos módulos.

Actualmente se ha iniciado el desarrollo del Módulo 5 que consiste en gráficos de control multivariante no paramétricos que se basan en el concepto de profundidad de datos pero debido a que aún no está terminado no es presentado en este trabajo. Una vez terminado los módulos para implementar gráficos de control se continuará aumentando funciones para el ámbito de fiabilidad de sistemas desde un contexto paramétrico y no paramétrico.

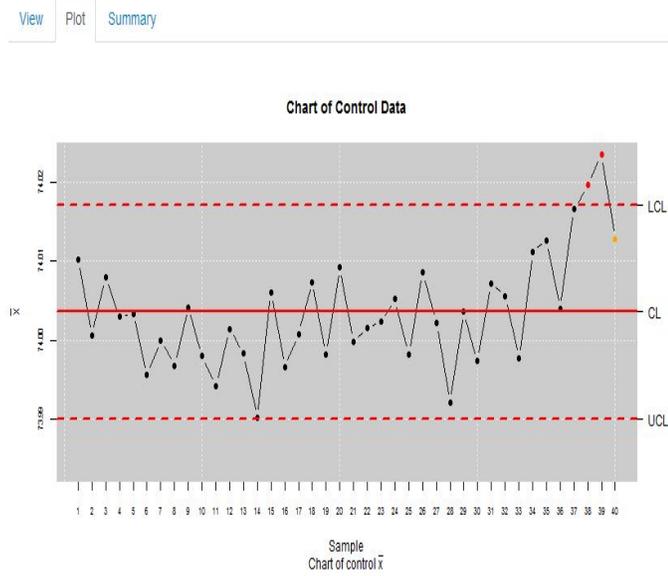


Figura 3.23: Gráfico de control xbar

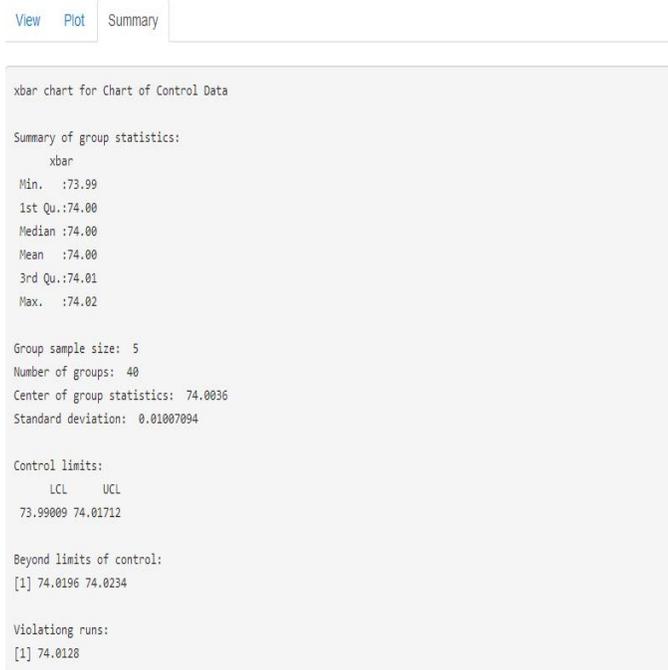


Figura 3.24: Estadísticos del gráfico xbar

Capítulo 4

Conclusiones y futuras líneas de desarrollo

La principal ventaja del uso de la interfaz gráfica para usuarios no iniciados en el lenguaje de programación R es que no trabajan directamente con el paquete `qcr` que está desarrollado en R sino que a través de selecciones en el panel pueden implementar cualquier gráfico de control de calidad. Desde el contexto de programación los objetivos que se han cumplido al desarrollar el paquete `qcr` y la interfaz gráfica son:

- Tener en un solo paquete integrados los gráficos de control de calidad más utilizados para el caso paramétrico y no paramétrico.
- Mejorar las salidas gráficas de los paquetes de control de calidad antes mencionados.
- Diseñar las funciones de tal forma que la interacción con el desarrollo de la interfaz gráfica sea más eficaz ya que la filosofía de desarrollo está en base al paquete **Shiny**.
- Facilitar la inclusión de nuevas funciones para la creación de gráficos de control ya que el diseño es modular, es decir, no depende de las anteriores implementaciones.
- No redundar en código ya que se utilizan las funciones implementadas en los paquetes `qcc` y `MSQC`.

- Permitir utilizar las funciones de otros paquetes en el caso de que ya estén implementadas para el cálculo de las medidas estadística y así solo se preocupa de las salidas gráficas.
- Implementar gráficos de control de calidad multivariante no paramétricas que no se encuentran actualmente desarrolladas en ningún paquete.

Sobre las futuras líneas de desarrollo se puede mencionar:

- En cada módulo en el panel de configuración se puede aumentar más parámetros del proceso, es decir se puede dejar que el usuario ingrese una media nominal y una varianza poblacional, así también parámetros de calibrado para la construcción de los límites de control. Por otro lado, dado el carácter modular del paquete y de la interfaz gráfica aumentar sus funcionalidades es muy fácil ya que ningún módulo es dependiente de otro módulo y por tanto no se requiere del conocimiento de como se encuentra desarrollado los otros módulos para crear nuevos módulos.
- Actualmente se ha iniciado el desarrollo del Módulo 5 que consiste en gráficos de control multivariante no paramétricos que se basan en el concepto de profundidad de datos pero debido a que aún no está terminado no es presentado en este trabajo. Una vez terminado los módulos para implementar gráficos de control se continuará aumentando funciones para el ámbito de fiabilidad de sistemas desde un contexto paramétrico y no paramétrico.

Apéndice A

Interfaz gráfica (Módulos)

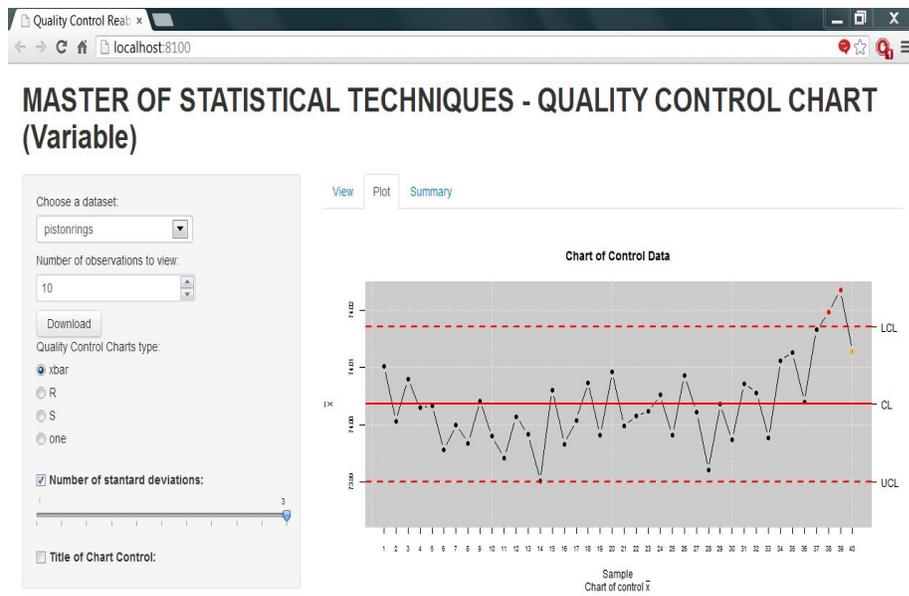


Figura A.1: Módulo para gráficos de control para variables cuantitativas

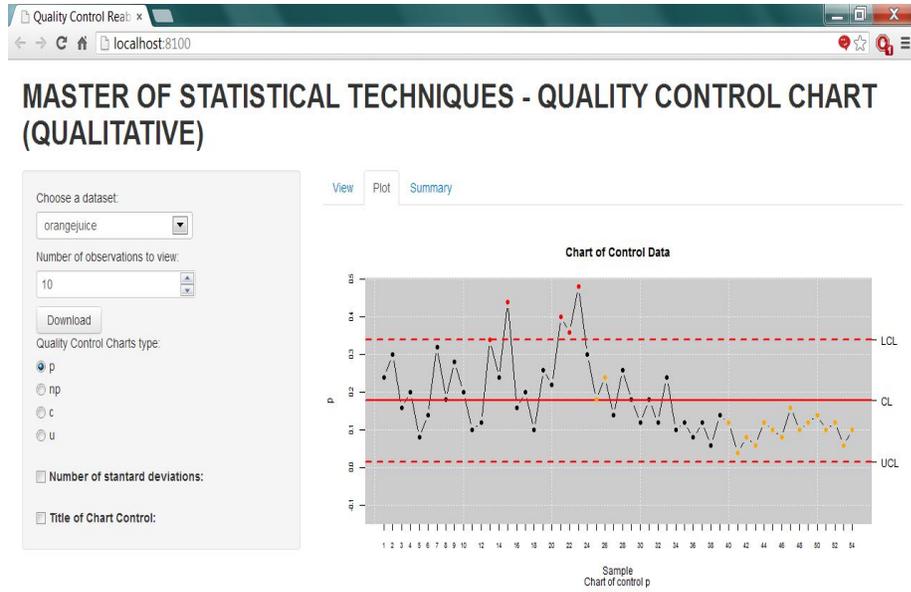


Figura A.2: Módulo para gráficos de control para variables cualitativas

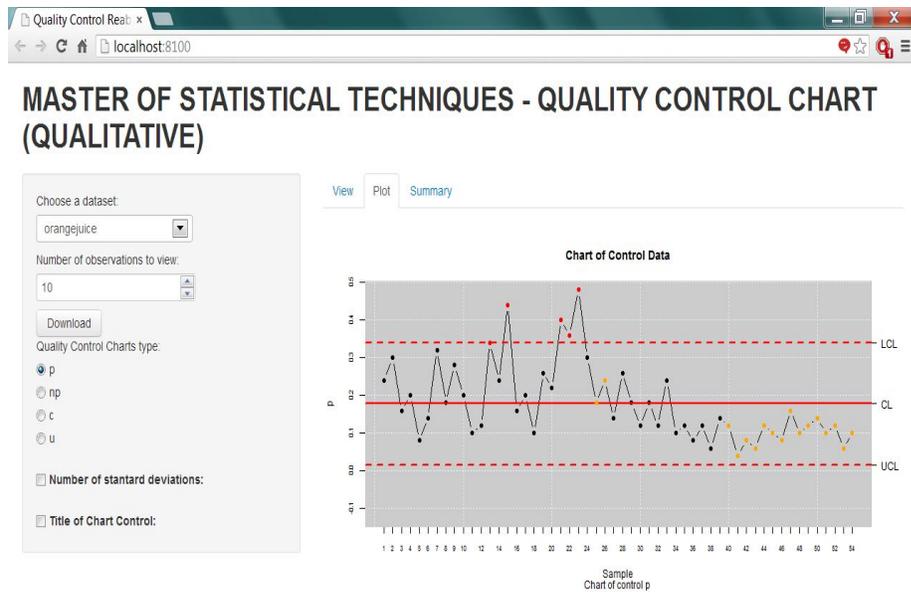


Figura A.3: Módulo para gráficos de control para variables dependientes

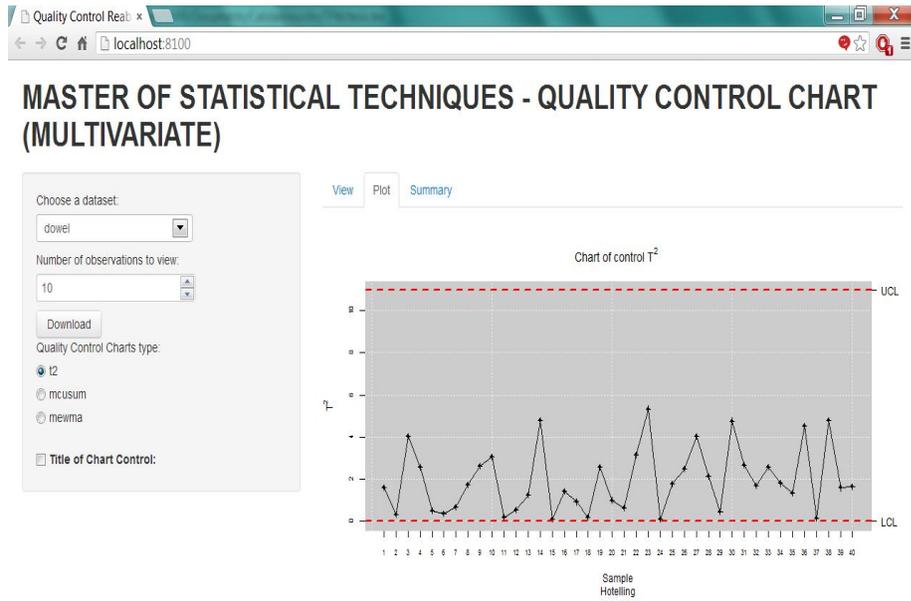


Figura A.4: Módulo para gráficos de control Mutivariante

Bibliografía

- 1 Abu-Shawiesh MO, Abdullah MB. New robust statistical process control chart for location. *Quality Engineering* 2000; 12: 149-159.
- 2 Albin S. L., Kang L., Sheha G., (1997), An and EWMA chart for individual observations. *Journal of Quality Technology*, 29, 1, 41-48.
- 3 Albing M. Process capability indices for Weibull distributions and upper specification limits. *Quality and Reliability Engineering International* 2009; 25: 317-334.
- 4 Alfaro Navarro J. L., Mtodos multivariantes en control estadstico de la calidad. Facultad de Ciencias Econmicas y Empresariales de Albacete. Universidad de Castilla-La Mancha.
- 5 Alt F. B., (1985), Multivariate Control Charts in *Encyclopedia of Statistical Sciences*, vol. 6. S. Kotz y N. L. Johson, Eds. Wiley, New York.
- 6 Alwan L. C., (1986), CUSUM Quality Control-Multivariate Approach, *Communications in Statistics*, 15, 3531-3543.
- 7 Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley.
- 8 Aparisi F., (1996), Hotellings T2 control chart with adaptive sample sizes, *International Journal of Production Research*, 34, 2853-2862.
- 9 Aparisi F., Garca Daz J. Carlos, Aumento de la potencia del grfico de control multivariante T2 de Hotelling utilizando seales adicionales de falta de control. *Estadstica espaola*. Vol 43, Nm. 148, 2001, pgs. 171 a 188.
- 10 Artiaga, R., Cao, R. y Naya, S. Local Polynomial estimation of TGA derivatives using logistic regression for pilot bandwidth selection. *Actas de la 5 Mediterranean Conference on Calorimetry and Thermal Analysis*, 86.

- 11 Artiaga, R., Garca, A., Varela, A., Mier, J. L., Naya, S. y Graa, M (2002). DMTA study of a nickel-titanium wire. *Journal of Thermal Analysis and Calorimetry*, 70, 199-207.
- 12 Bai, D. S. y Choi, I. S. (1995) X and R Control Charts for Skewed Populations. *Journal of Quality Technology*, 27, pp. 120-131.
- 13 Doganaksoy, N., Faltin, F. W. y Tucker. W.T. (1991), Identification of out-of-control characteristics in a multivariate manufacturing environment, *Communications in Statistics: Theory and Methods*, Vol. 20, pp. 2775-2790.
- 14 Blazek, L. W., Novic B. y Scott M. D. (1987): Displaying Multivariate Data Using Polyplots. *Journal of Quality Technolgy*. 19 (2) pp 69-74.
- 15 Borges W, Ho LL. A fraction defective based capability index. *Quality and Reliability Engineering International* 2001; 17: 447-458.
- 16 Brook, D, Evans D (1972), An approach to the probability distribution of cusum run length, *Biometrika* 59, 539-548.
- 17 Breyfogle, F.W. (2003). *Implementing Six Sigma*, second edition. New York: John Wiley and Sons.
- 18 Box G. E. P. y Luceo A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. Wiley. New York.
- 19 Capilla C, Ferrer A, Romero R. y Hualda A. (1999). Integration of statistical and engineering process control in a continuous polymerization process. *Technometrics*, 41, 1428.
- 20 Carot Alonso, V. (1996): *Control Estadstico de Calidad*. Servicio de Publicaciones de la Universidad Politecnica de Valencia, Valencia.
- 21 Castillo E. (2002). *Statistical Process Adjustment for Quality Control*. Wiley. New York.
- 22 Chakraborti, S., Van Der Laan, P. y Bakir, S. T. (2001) Nonparametric Control Charts: An Overview and Some Results. *Journal of Quality Technology*, 33, pp. 304-315.

- 23 Chandrasekaran, S. English, J. R. Disney R. L. (1995), Modeling and analysis of EWMA control schemes with variance-adjusted control limits, *IIE Transactions* 277, 282-290.
- 24 Chen KS, Pearn WL. An application of non normal capability indices. *Quality and Reliability Engineering International* 1997; 13: 355-360.
- 25 Chou YM, Owen DB. On the distribution of the estimated process capability indices. *Communications in Statistics. Theory and Methods* 1989; 18: 4549-4560.
- 26 Chua M. K., Montgomery D. C., Investigation and characterization of a control scheme for multivariate quality control. *Quality and Reliability Engineering International*, 1992, 8: 37-44. 260 BIBLIOGRAFA
- 27 Chung K., A Simplified Procedure for the Economic Design of Control Charts, *International Journal of Production Research*, Vol. 28, 1990, pp.1239-1246.
- 28 Crosier R. B. (1986), A new two-sided cumulative quality control scheme, *Technometrics* 28, 187-194. xewma.arl for zero-state ARL computation of EWMA control charts and xcusum.ad for the steady-state ARL.
- 29 Crosier R. B., (1988), Multivariate Generalization of Cumulative Sum Quality Control Schemes, *Technometrics*, 30, 291-303.
- 30 Crowder S. V. (1987), A simple method for studying run-length distributions of exponentially weighted moving average charts, *Technometrics* 29, 401-407.
- 31 Clements JA. Process capability indices for non normal calculations. *Quality Progress* 1989; 22: 49-55.
- 32 Dahlgaard, J.J. and Dahlgaard-Park, S.M. (1999). Lean production, TQM and economic development. *Proceeding of the Conference in Lean Thinking*, Stockholm, Sweden.
- 33 Deming, W.E. (1986), *Out of the Crisis*, Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA.
- 34 Deming, W.E. (1993), *The New Economics for Industry, Government, Education*, Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA.

- 35 Duncan A. J., The Economic Design of Charts when there is a multiplicity of Assignable Causes, *Journal of the American Statistical Association*, Vol. 66-333, 1971, pp. 107-121.
- 36 Duncan A. J., *Control de calidad y estadística industrial*. Ed. Alfaomega. México D.F. 1996.
- 37 Duncan A. J. (1986): *Quality Control and Industrial Statistics*, Homewood, IL: Irwin.
- 38 Eubank, R.L. (1988). *Spline Smoothing and Nonparametric Regression*. Marcel Dekker, New York.
- 39 Fuchs, C. y Kenett, R.S. (1998): *Multivariate Quality Control*. Marcel Dekker, New York.
- 40 George, S. (1992). *The Baldrige Quality System*, Wiley, New York, NY.
- 41 George, M. L. (2003). *Lean Sigma for service*. New York: McGraw-Hill.
- 42 Goel, A. L., Wu S. M. (1971), Determination of A.R.L. and a contour nomogram for CUSUM charts to control normal mean, *Technometrics* 13, 221-230.
- 43 Hoerl, R. W. (2001). Six Sigma Black Belts: what do they need to know? *Journal of Quality Technology* 33 4:391-435.
- 44 Hoerl, R. W., and R. D. Snee. (2002). *Statistical thinking: improving business performance*. Pacific Grove, CA: Duxbury Press.
- 45 Hotelling H., (1947), *Multivariate Quality Control*. *Techniques of Statistical Analysis*, (Eds. C. Eisenhart, M. Hastay y W. A. Wallis, McGraw-Hill), 111-184.
- 46 Ishikawa, K. (1994): *Introducción al Control de Calidad*. Daz de Santos.
- 47 Huber P. *Robust Statistics*. John Wiley and Sons: New York, 1981.
- 48 Jackson D. A., (1993), Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. *Ecology*, 74, 8, 2204-2214.
- 49 Jackson J. E. (1972). All Count Distributions Are Not Alike, *Journal of Quality Technology* 4(2):86-92.

- 50 Jackson J. E., (1985), Multivariate Quality Control, Communications in Statistics, 14 , pp. 2657-2688.
- 51 Jones, L. A. y Woodall, W. H. (1998) The Performance of Bootstrap Control Charts. Journal of Quality Technology, 30, No 4, pp. 362-375.
- 52 Juran, J.M. (1951). The Quality Control Handbook, 1 Ed., McGraw- Hill, New York, NY.
- 53 Juran, J. M. (1974). Juran's Quality Control Handbook, 3rd edition. McGraw-Hill, New York.
- 54 Juran, J.M. (1989). Juran on Leadership for Quality An Executive Handbook, The Free Press, New York, NY.
- 55 Kane VE. Process capability indices. Journal of Quality Technology 1986; 18: 41-52.
- 56 Kotz S, Johnson NL. Process Capability Indices. Chapman and Hall: London, 1993.
- 57 Kotz S, Johnson NL. Process capability indices-A review, 1992 2000 with discussion. Journal of Quality Technology 2002; 34: 2 - 53.
- 58 Kotz S, Lovelace CR. Introduction to Process Capability Indices: Theory and Practice. Arnold: London, 1998.
- 59 Kourti T. and MacGregor J. F., (1996), Multivariate SPC methods for process and product monitoring. Journal of Quality and Technology, 28: 409-428.
- 60 Lin PC, Pearn WL. Testing process capability for one-sided specification limit with application to the voltage level translator. Microelectronics Reliability 2002; 42: 1975-1983.
- 61 Liu R. 1995. Control Charts for Multivariate Processes. Journal of the American Statistical Association, 90 (432): 1389-1387.
- 62 Liu R, K Singh and J Tenkg. 2004. DDMA-charts: Nonparametric multivariate moving average control charts based on data depth. Allgemeines Statistisches Archiv, 88: 235-258.

- 63 Liu R. 1990. On a notion of data depth based on random simplices. *The Annals of Statistics*, 18 :405-414.
- 64 Liu R, J Parelius, K Singh, 1999. Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussions). *The Annals of Statistics*, 27: 783-858.
- 65 Lorenzen T. y Vance L., *The Economic Design of Control Charts: A Unified Approach*, Technometrics, Vol. 28, 1986, pp. 3-10.
- 66 Lowry C. A., Montgomery D. C., A review of multivariate control charts. *IIE Transactions* 1995, 27: 800-810.
- 67 Lowry, C. A. , Woodall, W. H. , Champ, C. W. and Rigdon, S. E (1992): A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics*, 34.
- 68 Lucas, J. M. Crosier R. B. (1982), Fast initial response for cusum qualitycontrol schemes: Give your cusum a headstart, *Technometrics* 24, 199- 205.
- 69 Lucas J. M.,Saccucci M. S. (1990), Exponentially weighted moving average control schemes: Properties and enhancements, *Technometrics* 32, 1-12.
- 70 Luceo. A. (1996). A process capability index with reliable confidence intervals. *Communications in Statistics - Simulation and Computation*. 25, 1, 235 245.
- 71 Mason N. Tracy J. and Young R, (1992), Multivariate control charts for individual observations. *Journal of quality technology*. 24, 2, 88-95.
- 72 Mason R. L., Tracy N. D., Young J. C., A practical approach for interpreting multivariate T2 control chart signals. *Journal of Quality Technology*, 1997, 29: 396-406.
- 73 Mason R. L., Tracy N. D., Young J. C., Decomposition of T2 for multivariate control chart interpretation. *Journal of Quality Technology* 1995 27: 99-108.
- 74 Mason R. L., Young J. C., Improving the sensitivity of the T2 statistic in multivariate process Control. *Journal of Quality Technology*, 1999, 31: 155-165.

- 75 Mc. Williams T. P., Economic Control Chart Designs and the In-Control Time Distribution: A Sensitive Study, *Journal of Quality Technology*, Vol. 21-2, 1989, pp. 103- 110.
- 76 Mitra, A. (1998): *Fundamentals of Quality Control and Improvement*. Prentice Hall.
- 77 Montgomery, D. C. (2009). *Introduction to Statistical Quality Control*. Sixth edition. John Wiley and Sons, Hoboken, NJ.
- 78 Nelson LS (1994). A Control Chart for Parts-Per-Million Nonconforming Items, *Journal of Quality Technology* 26(3):239-240.
- 79 Nembhard H, Changpetch P. (2006). Detected monitoring using cuscore charts for seasonal time series. *Quality and Reliability Engineering International*. 10,766-786.
- 80 Page E. S. (1954). Continuous inspection schemes. *Biometrika*, 41, 100- 15.
- 81 Page E. S., (1955), Control Charts with Warning Limits, *Biometrika*, 42, 243-257.
- 82 Page E. S., (1962). A modified Control Chart with warning lines, *Biometrika*, 49, 171-176.
- 83 Palmer K, Tsui K-L. A review and interpretations of process capability indices. *Annals of Operations Research* 1999; 87: 31-47.
- 84 Park, S.H. (2003), *Six Sigma for Quality and Productivity Promotion*, Asian Productivity Organization, Tokyo.
- 85 PearnWL, Chen KS. One-sided capability indices Cpu and Cpl: Decision making with sample information. *International Journal of Quality and Reliability Management* 2002; 19: 221-245.
- 86 Pearn W, Kotz S, Johnson NL. Distributional and inferential properties of process capability indices. *Journal of Quality Technology* 1992; 24: 216-231.
- 87 Pea D. (1987). *Estadística modelos y métodos. Modelos lineales y series temporales*. Alianza editorial. Madrid, 187-497.

- 88 Pignatello J. J. JR., Y Runger G. C., (1990), Comparisons of multivariate CUSUM charts, *Journal of Quality Technology*, 22, 173- 186.
- 89 Plackett R.L. y Burman, J.P. (1946). "The Design of Optimum Multifactorial Experiments", *Biometrika* 33 (4), 305-25.
- 90 Prasad S, Bramorski T. Robust process capability indices. *International Journal of Management Science* 1998; 26: 425-435.
- 91 Knoth S. (2003), EWMA schemes with non-homogeneous transition kernels, *Sequential Analysis* 22, 241-255.
- 92 Knoth S. (2004), Fast initial response features for EWMA Control Charts, forthcoming in *Statistical Papers*.
- 93 Rhoads, T. R., Montgomery D. C., Mastrangelo C. M. (1996), Fast initial response scheme for exponentially weighted moving average control chart, *Quality Engineering* 9, 317-327.
- 94 Ramsay, J. y Silverman, B. (1997). *Functional Data Analysis*. Springer- Verlag, New York.
- 95 Roberts S., (1959), Control Chart Tests Based on Geometric Moving Averages. *Technometrics* 1, 239-250.
- 96 Runger G. C., Alt F. B., Montgomery D. C., Contributors to a multivariate SPC chart signal. *Communications in Statistics - Theory and Methods*, 1996, 25: 2203-2213.
- 97 Ryan T. P., (2000). *Statistical Methods for Quality Improvement*; John Wiley: New York.
- 98 Seber G. A. F., (1984), *Multivariate Observations*, John Wiley and sons, Eds., New York.
- 99 Shewhart W. A., (1931). *Economic Control of Quality of Manufactured Products*, New York, D. Van Nostrand Co., Inc.
- 100 Shewhart, W. A. (1939): *Statistical Methods from the Viewpoint of Quality Control*. Graduate School of the Department of Agriculture, Washington, D.C.

- 101 Shewhart W.A., (1941), Contribution of Statistic to the Science of Engineering. Fluid Mechanics and Statistical Methods in Engineering. University of Pennsylvania Press, Philadelphia, PA, pp. 97-124.
- 102 Shewhart W.A., (1986): Statistical Methods from the Viewpoint of Quality Control. (Reprint). Dover Publications, New York.
- 103 Shewhart, Walter A. Economic control of quality of manufactured product. New York: D. Van Nostrand Company. pp. 501
- 104 Sparks R. S., Adolphson A., Phatak A., Multivariate process monitoring using the dynamic biplot. *International Statistical Review*, 1997, 65: 325- 349.
- 105 Steiner S. H. (1999), EWMA control charts with time-varying control limits and fast initial response, *Journal of Quality Technology* 31, 75-86.
- 106 Taguchi G., Introduction to Quality Engineering-Designing Quality into Products and Processes, 1986, Kraus International, Asian Productivity Organisation, Japan.
- 107 Timm N. H., Multivariate quality control using finite intersection tests. *Journal of Quality Technology*, 1996, 28: 233-243.
- 108 Tracy N. D., Young J. C. Y Mason R. L. (1992), Multivariate Control Charts for Individual Observations, *Journal of Quality Technology*, 24, 88-95.
- 109 Vermaat, M. B., Ion, R., Does, R.J. y Klaassen, C. (2003) A Comparison of Shewhart Individuals Control Chart Based on Normal, Non-Normal, and Extreme Value Theory. *Quality and Reliability Engineering International*, 19, pp. 337-353.
- 110 Chatfield, C., Collins, A. J. (1980). Introduction to multivariate analysis. Chapman and Hall.
- 111 Cuadras, C.M. (1991). Mtodos de Anlisis Multivariante. Editorial PPU.
- 112 Drapper, N. y Smith, H. (1998). Applied Regression Analysis. Wiley.
- 113 Fan, J. y Gijbels, I. (1996). Local Polynomial Modelling and its Applications. Chapman and Hall, London.

- 114 Goldstein M. y Dillon, W. R. (1984). *Multivariate Analysis: Methods and Applications*. Wiley.
- 115 Jambu, M. (1991). *Exploratory and Multivariate Data Analysis*. Boston Academic Press.
- 116 Jobson, J.D. (1992). *Applied Multivariate Data Analysis*. Springer- Verlag.
- 117 John Verzani (2005). *Using R for Introductory Statistics*. Chapman and Hall.
- 118 Mardia, K.V., Kent, J.T. y Bibby, J.M. (1994). *Multivariate Analysis*. Academic Press.
- 119 Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. John Wiley and Sons.
- 120 Pea Snchez de Rivera, D. (2002). *Regresin y Diseo de Experimentos*. Alianza Editorial.
- 121 Rencher, A.C. (1998). *Multivariate Statistical Inference and Applications*. Wiley.
- 122 Silverman, B. W. (1985) Some aspects of the spline smoothing approach to non-parametric curve fitting. *Journal of the Royal Statistical Society series B* 47, 152.
- 123 Venables, W. N. and Ripley, B. D. (1999) *Modern Applied Statistics with S-PLUS*. Third Edition. Springer.
- 124 Vilar Fernndez, J.M. (2003). *Modelos Estadsticos Aplicados*. Publicaciones de la UDC.
- 125 Somerville, S. E. and Montgomery, D. C. (1996). Process Capability Indices and Nonnormal Distributions. *Quality Engineering*, 9, 305-316.
- 126 Shore H. A new approach to analysing non-normal quality data with applications to process capability analysis. *International Journal of Production Research* 1998; 36: 1917-1933.
- 127 Tang LC, Than SE. Computing process capability indices for non normal data: A review and comparative study. *Quality and Reliability Engineering International* 1999; 15: 339-353.

- 128 US Military Handbook 217F, Reliability Prediction of Electronic Equipment. USA, 1991.
- 129 Vnnman, K. (1995). A Unified Approach to Capability Indices. *Statistica Sinica*, 5, 805-820.
- 130 Vnnman, K. (1998). Families of Capability Indices for One-sided Specification Limits. *Statistics*, 31, 43-66.
- 131 Vnnman, K. (2001). A Graphical Method to Control Process Capability. *Frontiers in Statistical Quality Control*, No 6, Editors: Lenz, H.-J. and Wilrich, P.-TH. Physica-Verlag, Heidelberg, 290-311.
- 132 Wasterhuis J. A., Gurden S. P, Smilde A. K., (2000), Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 51: 95-114.
- 133 WESTERN ELECTRIC, (1956), *Statistical Quality Control Handbook*, Select Code 700-444. Western Electric Corporation, Indianapolis, IN.
- 134 Wierda S. J., *Multivariate statistical process control-recent results and directions for future research*. *Statistica Neerlandica*, 1994, 48: 147-168.
- 135 Woodall WH, Ncube MM. *Multivariate CUSUM quality control procedures*. *Technometrics*, 1985, 27:285-292.
- 136 Woodall, W.H. (1985). The statistical design of quality control charts. *The Statistician*, 34. pp. 155-160.
- 137 Waldmann K. H. (1986), *Bounds for the distribution of the run length of one-sided and two-sided CUSUM quality control schemes*, *Technometrics* 28, 61-67.
- 138 Voehl, F. (2000). *Six sigma: a breakthrough strategy for results*. The STANDARD, Vol. 2000-1, The Newsletter of the Measurement Quality Division, American Society for Quality.
- 139 Welch, J. F. (2001). *Jack: straight from the gut*. New York: Warner Business Books.
- 140 Willeman, T. R. y Runger, G. C. (1996) *Designing Control Charts Using an Empirical Reference Distribution*. *Journal of Quality Technology*, 28, pp. 31-38.

- 141 Womack, J.P. and Jones, D.T. (1996). *Lean Thinking Banish Waste and Create Wealth in your Corporation*, Simon and Schuster, London.
- 142 Womack, J.P., Jones, D.T. and Roos, D. (1990). *The Machine that Changed the World*, Maxwell Macmillan International, New York, NY.
- 143 Woodall, W.H. (1985). The statistical design of quality control charts. *The Statistician*, 34. pp. 155-160.
- 144 Wu, H.-H. and Swain, J. J. (2001). A Monte Carlo Comparison of Capability Indices when Processes are Non-normally Distributed. *Quality and Reliability Engineering International*, 17, 219-231.
- 145 Zink, K. (1997), *Succesful TQM Inside Stories European Quality Award Winners*, Gower Publishing, London.
- 146 Zink, K. (1998), *Total Quality Management as a Holistic Management Concept. The European Model for Business Excellence*, Springer, Hamburg.