

# UNIVERSIDAD NACIONAL DE CÓRDOBA

Facultad de Ciencias Exactas, Físicas y Naturales

Tesis Doctoral



*Herramientas de monitoreo y control estadístico de atributos en procesos de alta calidad: estudio de resultados, implementación computacional y aplicación práctica.*

Autor: Mgter. Silvia Joekes

Director: Dr. Emanuel Pimentel Barbosa (hasta 15/11/2014)

Director: Dr. Smrekar Marcelo

Febrero 2016

*Herramientas de monitoreo y control estadístico de atributos en procesos de alta calidad: estudio de resultados, implementación computacional y aplicación práctica.*

por

Mgter. Silvia Joeques

Dr. Emanuel PIMENTEL BARBOSA

Dr. Marcelo SMREKAR

Directores

### **COMISIÓN ASESORA**

Dra. María Josefina POSITIERI

FCEFyN – UNC

Dr. Arnaldo MANGEAUD

FCEFyN – UNC

Esta Tesis fue enviada a la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de Córdoba para cumplimentar los requerimientos de obtención del grado académico de Doctor en Ciencias de la Ingeniería.

Córdoba, Argentina  
Febrero 2016

# TRIBUNAL

Dr. Jorge Domingo CUOZZO,  
Instituto Universitario Aeronáutico  
IUA

PhD Walter ROBLEDO  
Universidad Nacional de Catamarca  
UNDEC

Dr. Omar CHIOTTI,  
Ingeniería Argentina (INGAR)  
UTN - CONICET

*Dedicado a:*

*Inés y Emanuel*

## AGRADECIMIENTOS

Antes de introducir los contenidos formales de la presente investigación, quisiera agradecer a todas aquellas personas que colaboraron para la realización de esta tesis.

A mi hermana Inés que generó la adrenalina necesaria para comenzar este camino.

A mi primer Director, el Dr. Emanuel Pimentel Barbosa, quien me dio la oportunidad de desarrollar esta tesis confiando en mi y brindándome toda su experiencia y sabiduría.

Aunque ambos no están hoy físicamente conmigo para ver este proyecto terminado, estoy segura que estarían muy satisfechos con el resultado logrado.

A mi Director actual, el Dr. Marcelo Smrekar quien me brindó en todo momento su apoyo y sus conocimientos.

A mi comisión asesora, Dra. Positieri y Dr. Mangeaud quienes me acompañaron a lo largo de estos cinco años.

A los Directores de la carrera, Dr. Prato y Dr. Godoy quienes siempre estuvieron presentes cuando los necesité.

A mis evaluadores, los Doctores Cuzzo, Robledo y Chiotti, quienes se tomaron el paciente trabajo de leer mi tesis.

A la Universidad y particularmente a la Facultad de Ciencias Económicas que me posibilitó todos los medios para desarrollar este proyecto.

A Andrea, mi compañera de ruta, a Alejandra, secretaria del Doctorado por su paciencia y a Diana, que me consiguió todos los papers que necesité, hasta los imposibles.

Y por último, pero no por ello menos importante, a mi hermosísima familia.

## Tabla de contenidos

Capítulo 1: Introducción, objetivos y revisión de la literatura .....	1
1.1    Introducción .....	1
1.2    Motivación en el tema de tesis.....	2
1.3    Objetivos.....	4
1.4    Antecedentes.....	4
1.5    Organización de la tesis .....	5
Capítulo 2: Métodos tradicionales de Control Estadístico de Procesos (CEP) .....	8
2.1    Inicios del control estadístico de la calidad .....	8
2.2    Control Estadístico de Procesos.....	9
2.3    Determinación de los límites de control .....	12
2.4    Variables y atributos .....	12
2.5    Gráficos de control para atributos.....	13
2.5.1    Gráficos $p$ para fracción de no conformes.....	13
2.5.2    Gráfico $np$ para el número de unidades no conformes .....	15
2.5.3    Gráfico $c$ para el número de no conformidades.....	16
2.5.4    Gráfico $u$ para el promedio de no conformidades por unidad inspeccionada .....	17
2.6    Eficacia estadística de los gráficos de control .....	18
2.7    Algunos problemas relacionados con los gráficos de atributos .....	20
2.8    Necesidad de buscar nuevos procedimientos para el control de atributos .....	21
Capítulo 3: Métodos alternativos para gráficos de control en ausencia de normalidad .....	22
3.1    Introducción.....	22
3.2    Límites de probabilidad .....	23
3.3    Transformación de los datos .....	25
3.3.1    Gráfico $Q$ binomial.....	26
3.3.2    Arcoseno del gráfico $p$ .....	27
3.3.3    Bartlett, Anscombe, Freeman y Tukey.....	27
3.3.4    Gráfico $Q$ Poisson .....	28
3.3.5    Transformación raíz cuadrada .....	28
3.3.6    Transformación raíz cuadrada mejorada (ISRT).....	29
3.4    Transformación sobre los límites de control de los gráfico $p$ y $np$ .....	29
3.4.1    Límites de control con un ajuste en la expansión de Cornish-Fisher .....	30
3.4.2    Límites de control con dos ajustes en la expansión de Cornish-Fisher .....	34
3.4.3    Aplicación a un proceso de manufactura .....	38
3.5    Transformación sobre los límites de control de los gráficos $c$ y $u$ .....	41
3.5.1    Límites de control con un ajuste en la expansión de Cornish-Fisher .....	41
3.5.2    Límites de control con dos ajustes en la expansión de Cornish-Fisher .....	43
3.6    Comentario final .....	44

Capítulo 4: Gráfico de control para $p$ con muestreo doble.....	45
4.1    Introducción.....	45
4.2    Procedimiento del gráfico de control con muestreo doble.....	47
4.3    Medida de desempeño del gráfico.....	49
4.4    Ventaja del gráfico $np$ mejorado.....	51
4.5    Diseño y aplicación del muestreo doble al gráfico $np$ mejorado.....	52
4.6    Aplicación del MD a un proceso de pintura de autopartes.....	56
4.7    Comentario final.....	59
Capítulo 5: Gráficos de control basados en el conteo acumulado de unidades conformes.....	60
5.1    Introducción.....	60
5.2    Gráfico CCC para inspección secuencial.....	62
5.2.1  Fundamento teórico de los gráficos CCC.....	62
5.2.2  Límites de control.....	63
5.2.3  Aplicación del gráfico CCC para distintos valores de $p$ .....	64
5.2.4  Decisiones basadas en el gráfico CCC.....	66
5.2.5  Límites de control en función de la longitud promedio de corrida.....	66
5.3    El gráfico CCC- $r$ .....	68
5.3.1  Fundamento teórico de los gráficos CCC- $r$ .....	68
5.3.2  Límites de control.....	69
5.3.3  Límites de control en función de la longitud promedio de corrida.....	70
5.3.4  Fundamento teórico de las dos nuevas propuestas.....	73
5.3.5  Estudio de validación estadística de ambos procedimientos.....	75
5.3.6  Aplicación a un problema con datos de una planta de autopartes.....	77
5.4    Gráfico CCS para inspección por muestreo.....	79
5.5    Gráfico CQC para cantidades acumuladas.....	83
5.6    Comentario final.....	84
5.6.1  Comentario sobre otros estudios relacionados.....	85
Capítulo 6: Gráfico de control para procesos de atributos con exceso de ceros.....	88
6.1    Introducción.....	88
6.2    Modelos con exceso de ceros o modelos cero-inflados.....	91
6.3    Modelo lineal generalizado (MLG).....	93
6.3.1  Modelo lineal generalizado aplicado a la distribución binomial.....	94
6.3.2  Estimación de los parámetros.....	95
6.3.3  Modelo binomial con exceso de ceros.....	97
6.4    Selección del modelo adecuado.....	99
6.5    Determinación de los límites de control.....	100
6.6    Análisis de sensibilidad del gráfico de control.....	101
6.7    Aplicación del modelo ZIB a un proceso industrial.....	102
6.8    Comentario final.....	106

Capítulo 7: Conclusiones, aportes de la investigación y trabajos futuros .....	108
7.1    Resultados y Conclusiones .....	108
7.1.1  Desarrollo de un gráfico $p$ mejorado .....	108
7.1.2  Construcción de una regla empírica para determinar el gráfico $p$ adecuado.....	110
7.1.3  Extensión de los gráficos de control con muestreo doble .....	111
7.1.4  Estudio de los gráficos de control para procesos cercanos a cero defectos.....	112
7.1.5  Desarrollo de una nueva metodología para gráficos de control .....	114
7.2    Aportes originales de esta investigación.....	115
7.3    Trabajos futuros .....	116
Anexo I: Expansión de Cornish-Fisher .....	117
I.1    Método de expansión de Cornish-Fisher .....	115
I.2    La expansión de Cornish-Fisher aplicad al gráfico $p$ .....	115
I.2.1  Momentos (naturales y centrados) de la distribución binomial.....	118
I.2.2  Cumulantes.....	119
Anexo II: Publicaciones y Presentaciones en Congresos:.....	120
II.1   Publicaciones .....	115
II.2   Presentaciones en Congresos .....	115
Anexo III: Programas en R.....	122
Referencias bibliográficas .....	131

## Índice de Figuras

Figura 2.1: Forma básica del gráfico de control estadístico de calidad.....	10
Figura 2.2: Relación entre límites de tolerancia natural, límites de control y límites de especificación. (Montgomery, 1991). .....	11
Figura 3.1: Límites de probabilidad usados cuando la característica no es normal. ....	23
Figura 3.2: Riesgo $\alpha$ bilateral para aproximación normal versus Cornish-Fisher (un ajuste).....	33
Figura 3.3: Riesgo $\alpha$ unilateral para aproximación normal versus Cornish-Fisher (un ajuste).....	33
Figura 3.4: Riesgo $\alpha$ bilateral para aproximación normal versus Cornish-Fisher (dos ajustes).....	35
Figura 3.5: Riesgo $\alpha$ unilateral para aproximación normal versus Cornish-Fisher (dos ajustes).....	35
Figura 3.6: Riesgo alfa – Bilateral, ( $n = 20, p < 0,020$ ) .....	36
Figura 3.7: Riesgo alfa – Bilateral, ( $n = 10, p < 0,030$ ) .....	36
Figura 3.8: Riesgo alfa – Bilateral, ( $n = 5, p < 0,060$ ) .....	37
Figura 3.9: Gráfico $p$ de control bilateral, $n = 20, p = 0,015, \alpha = 0,0027$ .....	39
Figura 3.10: Gráfico $p$ de control bilateral, $n = 20, p = 0,004, \alpha = 0,0027$ .....	40
Figura 4.1: Procedimiento del gráfico de control con MD.....	48
Figura 4.2: Gráfico de control $np$ para los valores de la Tabla 4.6 .....	57
Figura 5.1: Gráfico CCC para tabla N° 5.1 .....	65
Figura 5.2: Gráfico CCC para tabla N° 5.2 .....	65
Figura 5.3: Gráfico CCC para tabla N° 5.3 .....	65
Figura 5.4: Curva ARL para el gráfico CCC-2 con máximo en $p_1 > p_0$ . ....	72
Figura 5.5: Curvas ARL para CCC-2 con ambos Métodos.....	76
Figura 5.6: Curvas ARL para CCC-3 con ambos Métodos.....	77
Figura 5.7: Curvas ARL para CCC-4 con ambos Métodos.....	77
Figura 5.8: Gráfico CCC-2 para el proceso de moldeo por inyección .....	79
Figura 6.1: Tipo de vehículo y localización de la pieza bajo estudio.....	103
Figura 6.2: Probabilidades observadas y estimadas para los modelos binomial y ZIB.....	105
Figura 6.3: Gráfico de control para la cantidad de tornillos con fisuras. ....	106

## Índice de Tablas

Tabla 3.1: Resumen de ecuaciones de transformación.....	28
Tabla 3.2: Valores mínimos de $p$ según tamaño de muestra y tipo de corrección .....	38
Tabla 3.3: Límites de control y riesgo $\alpha$ para el gráfico de la Figura 3.9.....	39
Tabla 3.4: Límites de Control y riesgo $\alpha$ para el gráfico de la Figura 3.10.....	40
Tabla 4.1: Diseño comparativo del gráfico $np$ con MS, con $ARL_{0min} = 370,4$ .....	52
Tabla 4.2: Diseño óptimo para gráficos $np$ con MS y MD con $ARL_{0min} = 370,4$ y $\gamma = 1,5$ .....	54
Tabla 4.3: Diseño óptimo para gráficos $np$ con MS y MD con $ARL_{0min} = 370,4$ y $\gamma = 2,0$ .....	54
Tabla 4.4: Diseño óptimo para gráficos $np$ con MS y MD con $ARL_{0min} = 200$ y $\gamma = 1,5$ .....	55
Tabla 4.5: Diseño óptimo para gráficos $np$ con MS y MD con $ARL_{0min} = 200$ y $\gamma = 2,0$ .....	55
Tabla 4.6: Muestreo doble para el control de piezas no conformes .....	58
Tabla 5.1: Conteo de conformes entre dos no conformes .....	64
Tabla 5.2: Conteo de conformes entre dos no conformes .....	65
Tabla 5.3: Conteo de conformes entre dos no conformes .....	65
Tabla 5.4: Valores del ARL para gráficos CCC- $r$ con $r = 2, 3$ y $4$ , $p_0 = 500$ ppm y $\alpha = 0,0027$ .....	71
Tabla 5.5: Curvas comparativas de ARL para los gráficos CCC- $r$ ( $\alpha = 0,0027$ ) .....	76
Tabla 5.6: Conteo de conformes hasta obtener 2 no conformes.....	78
Tabla 6.1: Valores de ARL para $p_0 = 0,0100$ , $\phi_0 = 0,5$ y $\alpha = 0,0027$ .....	102
Tabla 6.2: Tabla de frecuencia .....	104
Tabla 6.3: Elección del modelo y estimación de los parámetros .....	104
Tabla 6.4: Probabilidades estimadas para los modelos binomial y ZIB.....	105
Tabla 6.5: Valor esperado y variabilidad para los modelos binomial y ZIB.....	106

## GLOSARIO

- ACP: Probabilidad media de cobertura (Average Coverage Probability).
- ARL: Longitud promedio de corrida (Average Run Length).
- ARL<sub>0</sub>: Longitud promedio de corrida cuando el proceso está bajo control.
- ARL<sub>0min</sub>: Mínimo valor del ARL<sub>0</sub> (bajo control).
- ARL<sub>1</sub>: Longitud promedio de corrida cuando el proceso está fuera de control.
- ARL<sub>U</sub>: Longitud promedio de corrida corregida.
- ATS: Tiempo promedio hasta la señal (Average time to signal).
- $\alpha$  : Probabilidad de Error de tipo I.
- $\alpha_0$ : Valor prefijado para el error de tipo I, considerado usualmente como 0,0027.
- $\beta$  : Probabilidad de Error de tipo II.
- C<sup>4</sup>: Gráfico de conteo acumulado de conformidades en grupo.
- CCC: Conteo acumulado de conformes (Cumulative Count of Conforming).
- CCC<sub>G</sub>: Conteo acumulado de grupos de conformes (Cumulative Count of Conforming Groups)
- CCC<sub>1+γ</sub>: Gráfico CCC en dos etapas.
- CCC-*r*: Conteo acumulado de conformes hasta observar *r* no conformes.
- CCS: Conteo acumulado de muestras conformes (Cumulative Conforming Sample)
- CEP: Control Estadístico de Procesos.
- CQC: Conteo acumulado de cantidades (Cumulative Quantity Control Charts).
- CQC-*r*: Conteo acumulado de cantidades hasta observar *r* defectos.
- CUSUM: Suma acumulada (Cumulative Sum).
- EWMA: Promedio móvil exponencialmente ponderado (Exponentially Weighted Moving Average).
- LA: Limite de advertencia.
- LC: Línea central.
- LIC: Límite inferior de control.
- LSC: Límite superior de control.
- LSC<sub>1</sub>: Límite de control de la primera etapa de muestreo.
- LSC<sub>2</sub>: Límite de control para la segunda etapa de muestreo.
- LIE: Límite inferior de especificación.
- LSE: Límite superior de especificación.
- $\hat{\lambda}$ : Promedio de no conformidades por unidad en la muestra.
- $\lambda$ : Promedio de no conformidades por unidad en la población.
- $\lambda_0$ : Promedio de no conformidades por unidad cuando el proceso está bajo control.
- $\lambda_1$ : Promedio de no conformidades por unidad cuando el proceso está fuera de control.
- MLG: Modelo lineal Generalizado.

$n$ : Tamaño de la muestra.

$n_1$ : Tamaño de la primera muestra.

$n_2$ : Tamaño de la segunda muestra.

$n\hat{p}$ : Cantidad de unidades no conformes en la muestra.

$np$ : Cantidad de unidades no conformes en la población.

OC: Curva operatoria característica

$\hat{p}$ : Proporción de no conformes en la muestra.

$p$ : Proporción de no conformes en la población.

$p_0$ : Proporción de no conformes en la muestra cuando el proceso está bajo control.

$p_1$ : Proporción de no conformes en la muestra cuando el proceso está fuera de control.

SQC: Control estadístico de calidad (Statistical Quality Control).

TPM: Tamaño promedio de muestra.

$TPM_0$ : Tamaño promedio de muestra para un proceso bajo control.

$TPM_1$ : Tamaño promedio de muestra para un proceso fuera de control.

$\hat{u}$ : Promedio de no conformidades por unidad de inspección en la muestra.

$u$ : Promedio de no conformidades por unidad de inspección en la población.

$u_0$ : Promedio de no conformidades por unidad de inspección cuando el proceso está bajo control.

$u_1$ : Promedio de no conformidades por unidad de inspección cuando el proceso está fuera de control.

VSI: Intervalo variable de muestreo (Variable Sample Interval).

ZIB: Binomial con exceso de ceros (Zero-inflated binomial).

ZIP: Poisson con exceso de ceros (Zero-inflated Poisson).

## Resumen

Los procesos actuales, fruto de la evolución tecnológica, son procesos que se caracterizan por presentar una cantidad muy baja de unidades no conformes. En estas circunstancias, los procedimientos tradicionales de control ya no pueden ser empleados debido a que incrementan de manera significativa el riesgo de falsas alarmas. Esta situación motiva la necesidad de desarrollar una serie de alternativas más potentes que las clásicas para el control de procesos de atributos de alta calidad.

Al respecto, cuatro fueron los aspectos abordados en esta investigación. El primero consistió en la modificación de los límites de control del gráfico  $p$  tradicional mediante la incorporación de dos términos de ajuste. Los resultados mostraron que con estos nuevos límites de control se logra una mejora importante en el sentido de que posibilitan el monitoreo de procesos con una fracción de no conformidades  $p$  más pequeña. Además se pudo establecer una regla a efectos de considerar la elección apropiada de los límites de control para el gráfico  $np$ . El segundo consistió en el desarrollo del gráfico  $np$  de control con muestreo doble. Los resultados permitieron establecer que la eficacia de este procedimiento se basa principalmente en el hecho de que permite observar cambios pequeños o moderados en la proporción de unidades no conformes del proceso, sin aumentar el muestreo. El tercer aspecto abordado consistió en la revisión de los gráficos de control del tipo CCC, estudiando sus ventajas y limitaciones. En particular, debido a que en la literatura se presentan dos propuestas para los gráficos CCC- $r$ , ambas fueron evaluadas en función del ARL a efectos de establecer la más adecuada. El cuarto tema abordado consistió en el estudio de los procesos con “inflación de ceros”. Al respecto se desarrolló una nueva metodología, basada en los modelos lineales generalizados, para implementar gráficos de control en procesos binomiales con exceso de ceros. Se establecieron los límites de control y se estudió la sensibilidad del gráfico para diferentes valores de los parámetros y distintos tamaños de muestra.

En todos los casos analizados, se presentaron estudios con base en cálculos teóricos, simulaciones y análisis estadísticos ilustrados mediante aplicaciones prácticas con datos de procesos productivos industriales.

## Abstract

Today's high quality industrial processes are characterized by the presentation of a low fraction of non-conforming units. Under these circumstances, traditional control procedures can no longer be used, since they increase the risk of false alarms. Out of this situation, a need arises to develop a series of more powerful alternatives to control the processes of high-quality attributes.

Four aspects were addressed in this research. The first one consisted in a modification of the control limits of the traditional  $p$  chart by adding two terms of adjustment. The results showed an important improvement, since these new control limits allow for the monitoring of processes with a smaller non-conforming  $p$  fraction. Moreover, a rule that allows the proper selection of the control limits for the  $np$  chart could be established.

The second aspect consisted in the development of the double sample  $np$  control chart. The results showed that the efficacy of the procedure is mainly based on the fact that it allows for the observation of small or moderate changes in the proportion of non-conforming units without altering the sample.

The third aspect consisted in the study of the advantages and limitations of the CCC control charts. Since there are two proposals for CCC- $r$  charts in the literature, both of them were assessed regarding the ARL function, so as to establish the more adequate one.

The fourth and last aspect consisted in the study of the zero-inflated processes. A new methodology, based on the generalized lineal models, was developed in order to implement control charts in zero-inflated binomial processes. The control limits were established and the chart sensitivity for different parameters and different sample sizes was studied.

In all the cases analysed, we presented studies based on theoretical *calculus* simulations and illustrated statistical analyses through practical applications of illustrated data of industrial and productive processes.

## Resumo

Os atuais processos, fruto da evolução tecnológica, são processos que se caracterizam por apresentar uma quantidade muito baixa de unidades não conformes. Nestas circunstâncias, os procedimentos tradicionais de controle não podem ser empregados desde que acrescentam de maneira significativa o risco de falsos alarmes. Essa situação motiva a necessidade de desenvolver uma série de alternativas mais potentes que as clássicas para o controle de processos de atributos de alta qualidade.

A respeito disso, quatro deles foram abordados nesta pesquisa. O primeiro consistiu na modificação dos limites de controle do gráfico  $p$  tradicional mediante a incorporação de dois termos de ajuste. Os resultados mostraram que com estes novos limites de controle consegue-se uma importante melhora no sentido de que possibilitam o monitoramento de processos com uma fração de não conformidades  $p$  menor. Além disso, foi possível estabelecer uma regra aos efeitos de considerar a eleição apropriada dos limites de controle para o gráfico  $np$ . O segundo consistiu no desenvolvimento do gráfico  $np$  de controle com amostragem dupla. Os resultados permitiram estabelecer que a eficácia deste procedimento baseia-se principalmente no fato de permitir observar mudanças pequenas ou moderadas na proporção de unidades não conformes do processo, sem acrescentar a amostragem. O terceiro aspecto abordado consistiu na revisão dos gráficos de controle do tipo CCC, estudando suas vantagens e limitações. Em particular, devido a que na literatura apresentam-se duas propostas para os gráficos CCC- $r$ , ambas foram avaliadas em função do ARL aos efeitos de estabelecer a mais adequada. O quarto tema abordado consistiu no estudo dos processos com “inflação de zeros”. A respeito desenvolveu-se uma nova metodologia, baseada nos modelos lineares generalizados, para implementar gráficos de controle em processos binomiais com excesso de zeros. Estabeleceram-se os limites de controle e estudou-se a sensibilidade do gráfico para diferentes valores dos parâmetros e distintos tamanhos de amostra.

Em todos os casos analisados, apresentaram-se estudos com base em cálculos teóricos, simulações e análises estatísticas ilustradas por meio de aplicações práticas com dados de processos produtivos industriais.

## Capítulo 1

### Introducción, objetivos y revisión de la literatura

#### 1.1 Introducción

La industria mundial enfrenta, desde hace algunos años, una nueva realidad como consecuencia de los avances logrados principalmente por las revolucionarias leyes de oferta y demanda globalmente distribuidas. Por este motivo, en los últimos años, se ha observado en el ámbito de las empresas, tanto de producción como de servicios, un creciente énfasis en la calidad, enmarcada en la puesta en marcha de un proceso de mejora continua. El fundamento de este proceso se orienta hacia la potenciación de los recursos humanos y el uso de un enfoque racional y científico en el análisis de los problemas de manera de aprovechar toda la información que se genera en una organización. En este contexto, la utilización de los métodos estadísticos desempeña un papel de gran importancia y, en particular, el Control Estadístico de Procesos (CEP) constituye una de las herramientas básicas para llevar a la práctica la denominada filosofía de la Calidad Total.

Esta investigación trata del estudio, mejoramiento, ejecución y aplicación de modernos métodos estadísticos para el seguimiento y control de atributos críticos en procesos de alta calidad (“high quality processes”). Como resultado de la evolución tecnológica, estos procesos se caracterizan por presentar una “fracción no conforme”  $\hat{p}$ , de tipo binomial, con probabilidad  $p$  muy baja, o bien, con un “promedio de no conformidades por unidad”  $\hat{u}$ , de tipo Poisson, con un promedio  $u$  muy pequeño.

En estas circunstancias, los tradicionales gráficos de control de Shewhart, basados en la aproximación normal, y muy utilizados en la práctica, dejan de ser apropiados debido a que incrementan de manera significativa el riesgo de falsas alarmas. Además, presentan dificultades en la especificación del límite inferior de control, limitando la detección de posibles mejoras del proceso.

Para tratar con estas dificultades, se consideraron tres abordajes estadísticos diferentes los cuales fueron evaluados comparativamente mediante medidas de su desempeño (riesgos alfa, beta y ARL -“Average Run Length”). Ellos son: i) el perfeccionamiento del cálculo de los límites de control a través de la expansión de Cornish-Fisher; ii) la sustitución de las estadísticas  $\hat{p}$  o  $\hat{u}$  por una estadística G (número de conformidades entre las no conformidades, tipo geométrica o binomial negativa) y; iii) la utilización de versiones “con inflación de ceros” para  $\hat{p}$  y  $\hat{u}$ .

Los resultados de estudios, con base en cálculos teóricos, simulaciones y análisis estadístico se ilustran mediante aplicaciones prácticas con datos de procesos productivos industriales.

## **1.2 Motivación en el tema de tesis**

El Control Estadístico de Procesos tiene como finalidad minimizar la producción defectuosa, mantener la mejora continua del proceso y comparar la producción respecto de las especificaciones. Para ello, se utilizan gráficos que funcionan como sistemas de observación permanente de los procesos. Estos gráficos monitorean una o varias características del producto, proceso o servicio que se consideran relevantes, los cuales deben responder a ciertas consignas pre-establecidas, para aceptar que todo el sistema está bajo control, funcionando de manera estable. Si por el contrario, se observa un alejamiento de tales pautas, se debe advertir sobre una posible anomalía, buscando la causa que la produjo a fin de tomar acciones correctivas.

Desde el año 1920 en que Shewhart comenzó a desarrollar estas técnicas, han sido propuestos y utilizados con éxito una gran variedad de gráficos de control. Sin embargo, muchas características de los procesos actuales exigen un alto estándar de calidad, lo que hace necesaria la búsqueda de nuevas propuestas para este tipo de procesos.

Estas circunstancias dan fundamento a la presente tesis en la cual se presentan y analizan una serie de alternativas más potentes que las clásicas, para el control de procesos por atributos cuando la proporción de defectos, en condiciones normales, es muy baja.

El monitoreo y control de determinadas características o atributos críticos de calidad de un proceso se realizan sobre una proporción  $p$  (parámetro de proceso), que a menudo se controla a lo largo del tiempo mediante una estadística como  $\hat{p}$  (proporción muestral) o proporción “no conforme” con distribución binomial. Otras veces se realiza mediante la estadística  $G$ , referida al conteo de unidades conformes entre no conformes, con distribución geométrica, conocida también como CCC (“Cumulative Count of Conforming”). Además de este caso dicotómico, en el cual una unidad o producto producido se lo clasifica en conforme o no conforme mediante un proceso independiente de pruebas de Bernoulli, es también muy común en el caso de productos más elaborados, inspeccionar la calidad a través de un recuento de  $\hat{u}$  (número promedio de las no conformidades por “unidad de inspección”), que en condiciones de independencia estadística, sigue un modelo de Poisson con promedio  $u$  (o  $\lambda$ ).

Bajo ciertas condiciones especiales (como en el caso de  $np$  grande, digamos mayor a 10, donde  $n$  es el tamaño de la muestra, o el caso  $\lambda$  superior a 15, por ejemplo), las estadísticas  $\hat{p}$  y  $\hat{u}$  tienen generalmente sus distribuciones aproximadas por una distribución normal y el problema para

probar/controlar la estabilidad de  $p$  o de  $u$  se torna simple, conduciendo a los tradicionales gráficos  $p$  y  $u$  de Shewhart.

Sin embargo, en una situación más realista, como puede ocurrir y ocurre no sólo en el contexto de la calidad y la mejora de los procesos industriales, sino también en otros ámbitos de aplicación del control de calidad (por ejemplo, en procesos relacionados con la salud o en control ambiental, entre otros), en general  $n$  no es muy grande y  $p$  (probabilidad de defecto o no conformidad u otro evento adverso) es pequeña, algunas veces muy cercana a cero. Esta situación da lugar a los llamados "procesos de alta calidad", donde la frecuencia de ceros puede ser muy alta ("Zero-Inflated Process"), o donde el promedio  $u$  de la estadística  $\hat{u}$  es muy pequeña.

Cuando el valor esperado de no conformidades es pequeño (o muy pequeño), la aproximación normal ya no es adecuada, y deben ser considerados límites de control (cuantiles) "exactos" o buenas aproximaciones para la estadística que monitorea el parámetro  $p$ , o el parámetro  $u$ , lo que conduce a un tratamiento diferente de los datos (Montgomery, D .C., 2008, DeVor, *et al.* 2007).

Para tratar esta problemática, es posible encontrar cuatro abordajes en la literatura (como se detalla en el punto 1.4 Antecedentes):

1.a- Perfeccionamiento o mejora en el cálculo de cuantiles (límites de control) mediante la aplicación de transformaciones en  $\hat{p}$  o  $\hat{u}$  (por ejemplo, transformación arcoseno, transformación raíz cuadrada, transformación cuantílica Q, y otras) que conduzcan a la distribución de la estadística a una mayor proximidad con la distribución normal.

O, como una variante más elaborada de este enfoque:

1.b- Refinamiento en el cálculo de los cuantiles de la estadística  $\hat{p}$  o  $\hat{u}$ , mediante la fórmula de Cornish-Fisher, basada en los cumulantes de la estadística.

2. Aplicación de gráficos de control con muestreo doble como una alternativa a los gráficos de control para atributos con muestreo simple.

3. Sustitución de la estadística  $\hat{p}$  binomial por la estadística G geométrica (o binomial-negativa) o de la estadística  $\hat{u}$  de Poisson por la estadística E exponencial, para el caso de considerar el conteo de conformes entre no conformes.

4. Adaptación del modelo binomial para  $\hat{p}$  (o Poisson para  $\hat{\lambda}$ ) en el caso de excesos de ceros, resultando en los modelos ZIB "Zero-Inflated Binomial" o ZIP "Zero-Inflated Poisson".

En esta investigación, estos enfoques son abordados, ampliados y analizados comparativamente a través de medidas de desempeño (como riesgos alfa y beta, entre otros), realizando una discusión detallada de las ventajas y desventajas de cada una de estas metodologías y estableciendo la más adecuada para cada situación práctica. Los resultados son además evaluados mediante su aplicación con datos simulados y obtenidos de procesos industriales. En todos los casos se ha trabajado con el software estadístico R, desarrollando programas propios y empleando paquetes existentes.

### 1.3 Objetivos

Dos son los objetivos principales de esta tesis:

- i) Investigar los procedimientos posibles más adecuados y eficientes para el monitoreo y control estadístico de atributos en procesos de alta calidad, en función de la magnitud de los parámetros, el diseño de la muestra y las medidas de desempeño sobre la base de un estudio comparativo de resultados.
- ii) Desarrollar medios propios de implementación computacional de estos procedimientos y aplicarlos a datos simulados y a datos provenientes de procesos industriales de interés práctico.

### 1.4 Antecedentes

Existe una amplia variedad de libros y artículos referidos a la construcción y aplicación de los gráficos  $p$  y  $u$  de Shewhart, que constituyen el punto de partida de este proyecto. Entre ellos se pueden citar libros como Montgomery, (2008) o de bibliografía equivalente como Alwan (2000) y DeVor *et al.* (2007).

Quesenberry (1995), Tsai *et al.* (2006) y Wang (2009), desarrollaron procedimientos para mejorar el cálculo de cuantiles (límites de control) mediante la aplicación de transformaciones en  $\hat{p}$  o  $\hat{u}$ . Entre ellas se pueden mencionar, por ejemplo, la transformación arcoseno, la transformación raíz cuadrada, la transformación cuantílica  $Q$ , que conducen a la distribución de la estadística a una mayor proximidad con la normal.

Winterbottom (1993), Chen (1998), Chen y Cheng (1998), Acosta-Mejía (1999), Cavalcanti y Cordero (2006), Wang (2009) y Lee y Lin (1992), trabajaron en el refinamiento del cálculo de los cuantiles de la estadística  $\hat{p}$  o  $\hat{u}$ , mediante la fórmula de Cornish-Fisher, basada en los cumulantes de la estadística. En esta investigación, estos resultados se extendieron incluyendo términos del orden  $n^{-3/2}$  en un gráfico  $p$  mejorado.

El procedimiento de muestreo doble para monitorear la proporción de unidades no conformes, ha sido tratado en la literatura para el caso de muestras grandes (que verifican normalidad). Entre ellos se destacan; Costa *et al.* (2007), Daudin (1992), Rodrigues *et al.* (2011), Wu *et al.* (2007). En esta investigación, se adaptó la bibliografía existente al caso de  $p$  chico y muestras pequeñas que es la situación más común en la industria.

Calvin (1983), Goh (1987), Lucas (1989), Bourke (1991, 2001), Kaminsky *et al.* (1992), Chan *et al.* (2002, 2009), Nelson (1994), Wu *et al.* (2001), Kuralmani *et al.* (2002), Chen y Cheng (2010), Chen (2009), Zhang *et al.* (2004), Zhang *et al.* (2006, 2012) y Xie *et al.* (2002), efectuaron estudios respecto a la sustitución de la estadística  $\hat{p}$  binomial por la estadística  $G$  geométrica (o binomial-negativa) o de la estadística  $\hat{u}$  de Poisson por la estadística  $E$  exponencial, sobre la base del conteo de conformes entre no conformes. En esta tesis se evaluó la base teórica de estos gráficos a efectos de establecer el procedimiento más adecuado para el monitoreo de procesos industriales.

Otro tema tratado en la tesis es el de los modelos con exceso de ceros que incluyen a los modelos zero-inflated Poisson (ZIP) y zero-inflated binomial (ZIB) como casos particulares. Estos modelos se obtienen a partir de los modelos existentes combinando adecuadamente el modelo tradicional con una singular distribución de ceros. Xie *et al.* (2001) desarrollaron un gráfico  $c$  para el modelo ZIP. Sim y Lim (2008), desarrollaron varias propuestas, basadas en intervalos de confianza, para la construcción de gráficos de control para datos con inflación de ceros, tanto para el caso Poisson como para el binomial. Fatai *et al.* (2010), investigaron un gráfico de control basado en la distribución ZIB truncada (TZIB). Katemee y Mayureesawan (2012a, 2012b), propusieron gráficos basados en la media, mediana y varianza de la distribución geométrica y en otra investigación aproximaron a la distribución ZIP mediante una distribución chi-cuadrada no central. Vieira *et al.* (2000), desarrollaron el modelo binomial con inflación de ceros, como un modelo mixto Bernoulli-binomial. En esta tesis se utilizó la metodología de los modelos lineales generalizados para determinar el modelo que mejor ajusta los datos. La bibliografía de referencia fue, Nelder y Wedderburn (1972), Hinde y Demetrio (1998) y McCullagh y Nelder (1989).

## **1.5 Organización de la tesis**

El contenido de la tesis está distribuido a lo largo de 7 Capítulos y 3 Anexos, de acuerdo al siguiente esquema:

En el Capítulo 1 se presenta una breve introducción al tema de la tesis, con motivación, objetivos y revisión de la literatura.

El Capítulo 2 hace referencia a los métodos tradicionales de Control Estadístico de Procesos (CEP) para atributos y su implementación. Se determina el procedimiento estadístico que subyace a la determinación de los límites de control, el concepto de variables y atributos y los diferentes tipos de gráficos de atributos existentes hasta el presente. Se determina además, la forma de establecer la eficacia estadística de los gráficos de control. Finalmente se adicionan algunos comentarios respecto a problemas relacionados con el uso de estos gráficos y la necesidad de buscar nuevos procedimientos de control.

En el Capítulo 3 se introducen los gráficos de control de atributos con corrección. Dentro de estos procedimientos se consideraron: el cálculo de límites de control exactos, diferentes transformaciones para normalizar los datos y procedimientos que permiten realizar directamente ajustes sobre los límites de control. Al respecto, se efectuó una extensión de los resultados obtenidos por Winterbottom (1993) incluyendo términos del orden  $n^{-3/2}$  en un gráfico  $p$  mejorado de manera de posibilitar la corrección no sólo del sesgo y la asimetría de la distribución sino también de la curtosis. Se presenta también una aplicación a un problema industrial.

El Capítulo 4 surge como una extensión del Capítulo anterior y consiste en el desarrollo y aplicación del muestreo doble (MD) para monitorear la proporción de unidades no conformes basada en la utilización del gráfico  $p$  mejorado en lugar del gráfico tradicional de Shewhart. Se presenta el procedimiento de construcción del gráfico, las medidas de desempeño y la aplicación del muestreo doble al gráfico  $np$  mejorado. Además se proporcionan tablas para facilitar la elección de los parámetros del MD en función de las características de cada proceso. Se ilustran los beneficios del MD para el seguimiento de procesos de alta calidad mediante su aplicación a un proceso de pintura de autopartes.

El Capítulo 5 trata sobre los gráficos de control clasificados en la categoría de gráficos de conformidades acumuladas. Estos gráficos se conocen como gráficos tipo CCC y entre ellos se destacan los gráficos CCC, CCC- $r$ , CCS y CQC. Se realiza una revisión teórica de los gráficos CCC, extendiendo luego los resultados a los gráficos CCC- $r$ . Posteriormente se presentan dos propuestas, una basada en la maximización del ARL y otra en un procedimiento iterativo para obtener un ARL cuasi maximal e insesgado, las cuales son comparadas evaluando su desempeño en función del ARL bajo control. Finalmente se realiza una aplicación del gráfico CCC- $r$  para datos de un proceso de producción de autopartes, con análisis y discusión de los resultados.

El Capítulo 6 hace referencia al estudio de los gráficos de control basados en procesos que se caracterizan por presentar una gran cantidad de muestras conformes. Estos procesos se conocen como procesos “Zero-Inflated” o procesos con “inflación de ceros” e incluyen a los modelos Zero-Inflated Poisson (ZIP) y Zero-Inflated Binomial (ZIB) como casos particulares. El objetivo de este

Capítulo consiste en desarrollar una nueva metodología que permita determinar si un proceso es binomial con un número elevado de ceros debido a una tasa de no conformidades muy chica, o si se tiene una masa de probabilidad asociada con la presencia de ceros que es superior a la esperada para una distribución binomial estándar. Para ello se introducen los modelos lineales generalizados (MLG) y su aplicación al modelo binomial y al modelo Bernoulli-binomial (cero-inflado). Posteriormente se determina la estimación de los parámetros y el criterio de selección del modelo más adecuado para un conjunto dado de datos. Se determina el cálculo de los límites de control del gráfico para el modelo con inflación de ceros y de su sensibilidad para diferentes valores de los parámetros y distintos tamaños de muestra. Finalmente se presenta la aplicación de esta metodología a un problema industrial, con análisis y discusión de los resultados.

En el Capítulo 7 se presentan los resultados y se detallan las principales conclusiones del trabajo. Además se establecen posibles líneas para trabajos futuros.

**Anexos:**

Anexo I: Método de expansión de Cornish-Fisher y su aplicación al gráfico  $p$ .

Anexo II: Publicaciones y Presentaciones en Congresos

Anexo III: Programas en R.

## Capítulo 2

### Métodos tradicionales de Control Estadístico de Procesos (CEP) para atributos y su implementación

#### 2.1 Inicios del control estadístico de la calidad

La estadística como ciencia se remonta sólo dos o tres siglos, con gran auge durante el siglo XX. Las primeras aplicaciones se realizaron en astronomía, física y en ciencias biológicas y sociales. Sin embargo, no fue hasta los años 1920 que la teoría estadística comenzó a ser aplicada al control de calidad de manera efectiva como resultado del desarrollo de la teoría del muestreo.

El primero en aplicar métodos estadísticos en el control de calidad fue Walter A. Shewhart, de la Bell Telephone Laboratories, quien emitió un documento el 16 de mayo 1924, en el que presentó un bosquejo de un moderno gráfico de control. Shewhart continuó mejorando y trabajando en este tema y en 1931 publicó un libro sobre el control estadístico de calidad, "*Economic Control of Quality of Manufactured Product*". Este libro fijó las bases para la aplicación de los métodos estadísticos en el control de procesos.

En 1925 Fisher y Shewhart desarrollaron importantes aplicaciones de la estadística al entonces llamado Control Estadístico de la Calidad. A partir del año 1950, Deming revolucionó la industria japonesa cambiando el concepto del Control de Calidad al de Mejoramiento de la Calidad como una actividad continua que nunca termina. Su metodología se encuentra en el libro "Out of the crisis" publicado en 1986.

Desde el año 1980 ha habido un profundo crecimiento en el uso de los métodos estadísticos para el mejoramiento de la calidad, motivado en parte, por la amplia pérdida de negocios y mercados sufrida por muchas empresas.

Actualmente, las necesidades del mercado y la evolución tecnológica, han motivado que se esté considerando el monitoreo y control de procesos que cuentan con altos estándares de calidad, es decir, procesos en los cuales la proporción de unidades no conformes, en condiciones normales, es muy baja.

## 2.2 Control Estadístico de Procesos

El Control Estadístico de Procesos (CEP) es una herramienta estadística que se utiliza para conseguir el producto adecuado desde el inicio del proceso. Con este procedimiento se abarcan tres objetivos: a) seguimiento y vigilancia del proceso, b) reducción de la variación y c) menor costo por unidad.

Para implementar el CEP, el primer paso consiste en el análisis descriptivo de la característica bajo estudio. La distribución de frecuencias por ejemplo, permite tener un primer conocimiento respecto a la forma de la distribución subyacente de la característica. También es un indicador de la capacidad del proceso para cumplir con determinados requerimientos. Si no se observan comportamientos anormales como presencia de valores atípicos, por ejemplo, en esta etapa es posible estimar los parámetros de la distribución.

De todos los procedimientos utilizados para analizar datos de procesos productivos (servicios) probablemente los más valiosos son los gráficos de control. La aplicación más difundida de los gráficos de control corresponde a los gráficos de control de Shewhart. Estos gráficos sirven para descubrir y diferenciar causas comunes de causas especiales de variación y proporcionan un medio de vigilancia y retroalimentación en los sistemas operativos de control. Cuando Shewhart desarrolló los gráficos de control, él buscaba una forma económica y sencilla de controlar un proceso, es decir, que sólo se tomaran acciones sobre el proceso cuando ocurriera una causa especial de variación.

Básicamente, un gráfico de control es una representación de una característica de calidad medida o calculada a partir de una muestra, repitiendo este muestreo en el tiempo. El gráfico tiene una línea central (LC) que representa el valor medio de la característica correspondiente en estado de control, es decir, cuando solamente causas naturales afectan al proceso. Junto con la línea central se colocan otras dos líneas horizontales que corresponden a los límites de control, uno superior (LSC) y otro inferior (LIC). Estos límites se fijan de manera que, si el proceso se encuentra bajo control, prácticamente todos los valores muestrales graficados se ubicarán entre ellos. Cuando esto sucede no es necesario tomar ninguna acción sobre el proceso. En caso contrario, si uno o más valores muestrales se ubican más allá de estos límites, el proceso estará fuera de control y serán necesarias acciones de investigación y corrección (Figura 2.1). Además, se espera que todas las muestras exhiban un patrón de comportamiento aleatorio dentro de estos límites. Si ello no ocurriera, se considerará que hay una causa especial de variación y, en consecuencia, el proceso estará fuera de control.

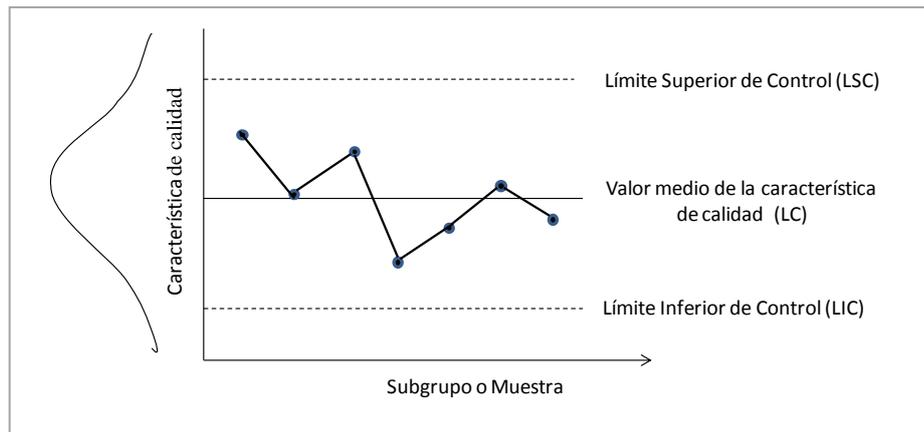


Figura 2.1: Forma básica del gráfico de control estadístico de calidad

Desde el punto de vista estadístico, un gráfico de control es básicamente una prueba de hipótesis de que el proceso está bajo control estadístico y como cualquier prueba de hipótesis, se pueden cometer los errores Tipo I (con probabilidad  $\alpha$ ) y Tipo II (con probabilidad  $\beta$ ) al tomar una decisión. Este concepto, netamente estadístico, muy pocas veces se conoce y en consecuencia se suele proceder de manera equivocada.

Para que un gráfico de control funcione adecuadamente, se debe establecer un plan de muestreo que minimice la variación del proceso dentro de cada muestra o subgrupo. Esto es así, debido a que se supone que las unidades producidas bajo condiciones similares de operación, constituyen un subgrupo homogéneo. Por lo tanto, la variación presente entre unidades de un mismo subgrupo, reflejan sólo la variación dentro del subgrupo. Esta variación se conoce habitualmente como variación natural del proceso, variación por causas comunes o variación de corto plazo y para los estadísticos es la medida del error experimental. Esta variación no está influenciada por el efecto de diferentes máquinas, diferentes operadores, cambios de materia prima, mantenimiento de equipos, temperatura ambiente o humedad, factores que suelen tener efecto entre subgrupos. La desviación estándar estimada sólo a partir de la variación dentro de los subgrupos, representa la consistencia de un proceso en un período corto de tiempo y por ello se la conoce como variación de corto plazo.

Las muestras o subgrupos se obtienen luego a lo largo del proceso a intervalos de tiempo regulares (frecuencia de muestreo), de manera que, de verificarse causas especiales, ellas se produzcan entre subgrupos. La desviación estándar estimada de la combinación de la variación dentro y entre subgrupos (variación total), refleja el desempeño del proceso sobre un largo período de tiempo.

Cuando un proceso se encuentra en “perfecto estado de control” la variación entre los promedios de subgrupos debe ser nula y, en consecuencia, las desviaciones de corto y largo plazo deberían ser iguales. Evidentemente, el estado de “perfecto control” rara vez se encuentra en la

práctica y la variación entre medias de subgrupos es generalmente mayor a la variación dentro de subgrupos o de corto plazo.

Si un proceso se encuentra “bajo control estadístico” el estudio de su capacidad se puede entender como un estudio de la predictibilidad de un proceso o de su habilidad para lograr determinados requerimientos.

Entonces, teniendo en cuenta todos estos conceptos, en una primera etapa de control, se identifican y eliminan las causas especiales de variación para estabilizar el proceso y en una segunda etapa se consolida la estabilidad mediante acciones de mejora y reacción a causas especiales. De esta manera se va realimentando el sistema

Es importante destacar que no existe ninguna relación entre los límites de control y los límites de especificación de un proceso. Los primeros surgen de la variabilidad natural del proceso medida por la desviación estándar de la característica bajo estudio, esto es, por los límites de tolerancia natural del proceso. Los límites de especificación, en cambio, son determinados por el departamento de diseño en base a requerimientos de clientes.

Debería conocerse la variabilidad inherente del proceso al fijar las especificaciones, pero no existe relación matemática o estadística alguna entre los límites de control y los límites de especificación. La relación entre estos límites se observa en la Figura 2.2.

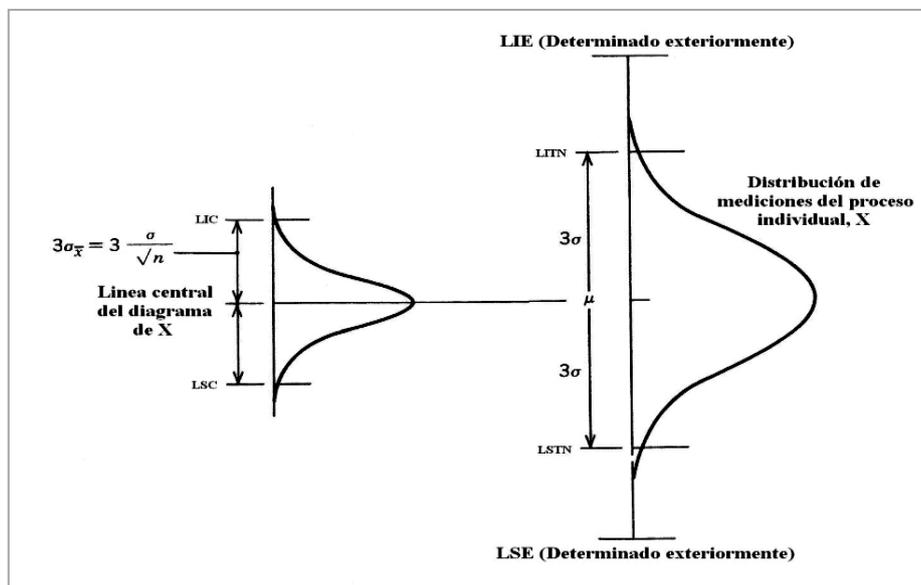


Figura 2.2: Relación entre límites de tolerancia natural, límites de control y límites de especificación. (Montgomery, 1991).

### 2.3 Determinación de los límites de control

La base estadística para definir los límites de control es común para todos los gráficos de Shewhart.

Si  $W$  es un estadístico que describe una determinada característica de calidad siendo  $\mu_w$  y  $\sigma_w^2$  su media y su varianza, los límites de control se definen como;

$$LIC = \mu_w - k\sigma_w ; \quad LC = \mu_w ; \quad LSC = \mu_w + k\sigma_w \quad 2.1$$

donde  $k$  es la distancia de los límites de control a la línea central expresada como un múltiplo de  $\sigma_w$ . Habitualmente se selecciona  $k = 3$ .

La elección de los límites de control es una de las tareas más críticas al momento de iniciar el control de procesos.

Si se eligen límites demasiado alejados de la línea central se corre el riesgo de tomar la decisión de que el proceso se encuentra bajo control cuando en realidad no lo está (error de tipo II, cuya probabilidad es  $\beta$ ).

Tomando límites demasiado cercanos a la línea central se corre el riesgo de decidir que el proceso está fuera de control cuando en realidad no lo está (error de tipo I, cuya probabilidad es  $\alpha$ ).

En general, si se supone que la característica se distribuye normalmente y se toman límites para  $k = 3$ , se encuentra a partir de la tabla normal estándar que,  $\alpha$ , la probabilidad de un error de tipo I es de 0,0027. Es decir, una falsa alarma se producirá solamente en 27 de 10.000 puntos muestrales. Por otro lado, la probabilidad de que un punto exceda a los límites de control en una sola dirección es de 0,00135.

Por lo general, se recomienda el uso de límites  $3\sigma$  porque dan buen resultado en la práctica pero en algunas situaciones especiales pueden emplearse otros valores de  $k$ .

### 2.4 Variables y atributos

Los gráficos de control se clasifican en dos tipos: gráficos para variables y gráficos para atributos.

Cuando la característica de calidad puede medirse y expresarse como un número se la llama variable. Ejemplos incluyen dimensiones tales como: el diámetro, la temperatura, el ancho y el volumen. Cuando una característica es variable, por lo general es necesario monitorear tanto el valor medio de la característica como su variabilidad. Para ello se utilizan los llamados *gráficos de control para variables*.

Sin embargo, existen muchas características cualitativas que no se miden en una escala cuantitativa. En estos casos, el interés consiste en determinar si una unidad acuerda o no acuerda con ciertas características, si posee ciertos atributos o si interesa contar el número de defectos que aparecen en cada unidad de producto. Los gráficos de control para estas características se denominan *gráficos de control para atributos*.

Como esta investigación trata del estudio de métodos estadísticos para el seguimiento y control de atributos críticos en procesos de alta calidad, se realizará el enfoque sobre estos últimos.

## 2.5 Gráficos de control para atributos

Algunas características de calidad no pueden ser representadas convenientemente por medio de variables cuantitativas. En estos casos, las unidades de producto se clasifican en “conformes” o en “no conformes” según la característica acuerde o no con las especificaciones. Las características de calidad de este tipo se denominan atributos. Los datos de tipo atributo toman solamente dos valores: conforme/no conforme, pasa/no pasa, funciona/no funciona, presente/ausente. También se consideran atributos aquellas características que se registran en términos de si/no, como por ejemplo la presencia de una enfermedad, la rugosidad de una placa cuya conformidad solo la medimos en términos de aceptable/no aceptable, las imperfecciones de pintura en una pieza, la presencia/ausencia de un sensor, etc.

Existen al momento, cuatro tipos de gráficos de control por atributos:

- ✓ Gráfico  $p$  para la fracción de no conformes
- ✓ Gráfico  $np$  para el número de unidades no conformes
- ✓ Gráfico  $c$  para el número de no conformidades por unidad
- ✓ Gráfico  $u$  para el promedio de no conformidades por unidad inspeccionada

### 2.5.1 Gráficos $p$ para fracción de no conformes

Este tipo de gráfico se utiliza cuando se tiene una característica de calidad bajo estudio que surge de clasificar a un producto en categorías generalmente relacionadas con la conformidad o no conformidad (disconformidad) respecto a las especificaciones.

La *fracción no conforme* se define como el cociente entre el número de unidades no conformes en la población y el número total de unidades que componen dicha población. Cada unidad de producto puede ser examinada respecto de una o varias características cualitativas. Si la unidad inspeccionada no acuerda con la especificación en una o más características, se la clasifica como no conforme o disconforme. Habitualmente, la fracción no conforme se expresa en forma decimal aunque puede también indicarse en porcentaje.

Cuando es posible suponer que un proceso funciona de manera estable, de manera que la probabilidad de que cualquier unidad deje de cumplir con las especificaciones (no conforme) es  $p$ , y que las unidades producidas en forma sucesiva son independientes, entonces cada unidad producida es una realización de una variable aleatoria Bernoulli con parámetro  $p$ .

Luego, si se selecciona una muestra aleatoria de  $n$  unidades y si  $X$  representa el número de unidades no conformes,  $X$  tiene distribución binomial con parámetros  $n$  y  $p$ .

Es decir:

$$P(X = x) = C_n^x p^x (1 - p)^{n-x} \quad , \quad x = 1, 2, \dots, n \quad 2.2$$

Cuando la característica de interés es la *fracción muestral de unidades no conformes* esta estadística se define como el cociente entre el número de unidades no conformes en la muestra  $X$ , y el tamaño de la muestra  $n$ .

En este caso,  $\hat{p} = X/n$  con media y varianza dadas por,  $\mu_p$  y  $\sigma_p = p(1 - p)/n$ .

Cuando se conoce la fracción  $p$  de unidades no conformes de un proceso de producción, los límites de control resultan:

$$LIC = p - 3 \sqrt{\frac{p(1 - p)}{n}} ; \quad LC = p ; \quad LSC = p + 3 \sqrt{\frac{p(1 - p)}{n}} \quad 2.3$$

El procedimiento de control consiste en tomar sucesivas muestras de  $n$  unidades, contar dentro de cada muestra el número de unidades no conformes llevando este valor al gráfico. Mientras  $\hat{p}$  permanezca dentro de los límites de control y la secuencia de puntos no señale ninguna pauta distinta a la que puede surgir por mero azar, es posible establecer que el proceso está bajo control a un nivel  $p$  de fracción no conforme. Si por el contrario, se observa algún punto fuera de los límites de control o un patrón inusual, se puede concluir que la fracción no conforme ha cambiado a un nivel diferente y que el proceso está fuera de control.

La evaluación del gráfico puede basarse en el error de tipo I, que es la probabilidad de que  $\hat{p}$  no se encuentre entre los límites superior e inferior del gráfico. Cuando  $p$  es conocido, generalmente se utiliza el gráfico con un error tipo I igual a 0,0027 y límites de control  $3\sigma$ .

Cuando se desconoce  $p$ , debe estimarse a partir de los datos. El procedimiento a seguir consiste en seleccionar  $m$  muestras preliminares, cada una de tamaño  $n$ . Como norma general,  $m$  estará comprendido entre 20 y 25. Si  $x_i$  es el número de unidades no conformes en la muestra  $i$ , la fracción no conforme en la muestra se calcula como  $\hat{p}_i = x_i/n$ ; para  $i = 1, 2, \dots, m$  y la media de estas fracciones  $\bar{p} = \sum_{i=1}^m \hat{p}_i/m$ , estimará la media del proceso, siendo ahora los límites de control dados por:

$$LIC = \bar{p} - 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}; \quad LC = \bar{p}; \quad LSC = \bar{p} + 3 \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} \quad 2.4$$

Estos límites de control se consideran como límites de prueba y sirven para determinar si el proceso estaba bajo control cuando las  $m$  muestras iniciales fueron seleccionadas. Si todos los puntos caen dentro de los límites de control y no se observa ninguna pauta anormal, se determina que el proceso se encontraba bajo control cuando se tomaron las  $m$  muestras y los límites de prueba serán válidos para controlar la producción actual y la futura. (Este procedimiento se conoce también como Fase I).

Como el gráfico de control se construye basado en la aproximación normal, cuando  $p$  es conocido, el error de tipo I no se aleja mucho del nivel nominal siempre que  $n$  sea grande (Teorema Central del Límite). Sin embargo, cuando  $p$  es desconocido, el verdadero error de tipo I no sólo depende de  $n$  sino también depende fuertemente del verdadero valor de  $p$ . Cuando el tamaño de la muestra es suficientemente grande, se espera que el gráfico permita controlar adecuadamente el proceso aún cuando la fracción de no conformes sea pequeña, como sucede en los procesos de alta calidad.

Actualmente hay muchos trabajos en los cuales se determina que el gráfico de control tradicional para  $p$  (que es el que se ha presentado), tiene serias dificultades para controlar adecuadamente un proceso cuando  $n$  y  $p$  son pequeños. Esto se debe a que el error de tipo I estará alejado del valor nominal 0,0027 y en consecuencia se producirán excesivas señales de falsa alarma.

También se utiliza el gráfico  $p$  en el caso de inspección 100%. Es común que al final de un día o un turno de trabajo se realice una inspección completa de las unidades producidas y en ese momento se grafique en un gráfico  $p$  la proporción de unidades no conformes. Como la cantidad de unidades producidas varía de turno en turno, de día en día, etc., los valores de  $n$  serán diferentes por lo que se necesitará implementar un gráfico de control con tamaño muestral variable.

### 2.5.2 Gráfico $np$ para el número de unidades no conformes

La variable *cantidad de unidades no conformes* es una variable aleatoria discreta porque puede tomar un número finito o infinito numerable de valores. Los gráficos  $np$  se utilizan para controlar el número de unidades no conformes en una muestra.

Considerando la población de unidades que podría producir el proceso trabajando siempre bajo las mismas condiciones, una cierta proporción  $p$  de éstas serán no conformes. Entonces, la probabilidad de tomar una unidad y que resulte no conforme es  $p$ . Luego, en una muestra de  $n$  unidades, la probabilidad de encontrar: 0, 1, 2...,  $n$  no conformes, está dada por una distribución

binomial con parámetros  $n$  y  $p$ . En esta situación, el promedio de la población es  $np$  y la varianza es  $np(1 - p)$ .

Los gráficos de control  $np$ , se construyen de la siguiente manera:

$$LIC = np - 3\sqrt{np(1 - p)} ; \quad LC = np ; \quad LSC = np + 3\sqrt{np(1 - p)} \quad 2.5$$

Cuando no se cuenta con un valor estándar para  $p$ , puede utilizarse  $\bar{p}$  para estimar a  $p$ . En este caso  $\bar{p}$  sustituirá al valor de  $p$  en los límites previos indicados (Ecuación 2.5).

Una vez obtenidos los límites de prueba, se grafica el número de unidades no conformes encontradas en las sucesivas muestras. Si no hay puntos fuera de los límites de control ni se encuentran patrones no aleatorios, se adoptan los límites calculados para controlar la producción futura.

Para las personas con poco entrenamiento estadístico, este gráfico suele ser más fácil de interpretar que el gráfico  $p$ .

### 2.5.3 Gráfico $c$ para el número de no conformidades

Una unidad no conforme es una unidad que presenta al menos una no conformidad. Sin embargo, dependiendo de su naturaleza o gravedad, es muy posible que una unidad contenga varias disconformidades. Por ejemplo, el interés puede estar en determinar el número de defectos funcionales en un dispositivo electrónico más que en determinar si el dispositivo funciona o no funciona. En estos casos puede ser más conveniente considerar el *número de no conformidad por unidad* en lugar de analizar la fracción de unidades no conformes.

Durante el proceso de producción, se retiran unidades a intervalos regulares y se cuenta el número total de no conformidades por unidad. En cada unidad se pueden encontrar, ninguna, una o más de una no conformidad. Los resultados obtenidos de este conteo, constituyen una variable aleatoria discreta con distribución de Poisson, dada por:

$$P(x) = \frac{e^{-c} c^x}{x!}, \quad x = 0, 1, 2, \dots \text{ y } c > 0 \quad 2.6$$

donde  $x$  es el número de no conformidades y  $c$  es el parámetro de la función. El promedio de la población es  $c$  y la varianza también es  $c$ .

Asumiendo que la distribución Poisson es adecuadamente aproximada por la distribución normal, generalmente cuando  $c \geq 5$  ( Xie, 2002), los límites de control  $3\sigma$  para el número de no conformidades por unidad de producto, están dados por:

$$LIC = c - 3\sqrt{c} ; \quad LC = c ; \quad LSC = c + 3\sqrt{c} \quad 2.7$$

Se pueden efectuar gráficos de control para el número total de no conformidades por unidad de producto o por unidad de inspección (varias unidades de producto). Estos gráficos de control se basan en la distribución Poisson que requiere que el número de oportunidades o lugares potenciales para las no conformidades sea muy grande y que la probabilidad de ocurrencia de la no conformidad en cualquier lugar sea muy pequeña y constante.

La unidad de inspección debe ser la misma en cada muestra. Es decir cada unidad de inspección debe tener siempre la misma probabilidad de que se produzcan las no conformidades.

Hay que tener en cuenta que la probabilidad de producir una falsa alarma por situarse el punto por encima del límite de control superior es diferente que la de situarse por debajo del límite inferior (colas superior e inferior diferentes) debido a que la distribución de Poisson es asimétrica.

Cuando no se conoce el parámetro  $c$ , se lo debe estimar a partir de una muestra preliminar de unidades de inspección. En este caso se obtiene el número promedio de no conformidades observadas en la muestra preliminar. El valor obtenido en la estimación  $\bar{c}$ , sustituirá al valor de  $c$  en los límites previos indicados en la Ecuación 2.7.

Es importante tener en cuenta que en muchas situaciones prácticas, los datos sobre las causas que producen las no conformidades aportan siempre mayor información que los relativos a la cantidad de unidades no conformes ya que habitualmente existen diversos tipos de no conformidades. En estos casos es necesario el desarrollo de planes de acción para condiciones fuera de control.

El gráfico de Pareto es muy importante para analizar la frecuencia asociada con cada tipo de no conformidad. Por lo general se observa que un pequeño número de no conformidades es causa de la mayor parte de los problemas y cuando es posible eliminar estas causas, se logrará una drástica mejora en la calidad.

Otra técnica útil es el diagrama de causa y efecto que permite determinar las posibles fuentes de no conformidades en los productos y sus interrelaciones. Además permite aumentar el nivel de conocimiento tecnológico del proceso.

#### **2.5.4 Gráfico $u$ para el promedio de no conformidades por unidad inspeccionada**

Supongamos que se desea controlar el número de no conformidades en un proceso y se define una unidad de inspección conformada por 10 unidades. En este caso, es posible trabajar con un gráfico  $c$  como fue dado anteriormente. Sin embargo, puede suceder que la variable que se desea

controlar sea el *promedio de no conformidades por cada unidad* de inspección en lugar del total de no conformidades para las 10 unidades (unidad de inspección).

Luego, si se encuentran  $x$  no conformidades totales en una muestra de  $n$  unidades de inspección, el número promedio de no conformidades por unidad de inspección es;  $u = X/n$  donde  $X$  es una variable aleatoria con distribución Poisson.

Como en el caso de los gráficos  $c$ , en una primera etapa se toman  $N$  unidades de inspección (más de 25 ó 30) a intervalos regulares. Se cuenta en cada unidad el número de no conformidades y se lo registra. Luego, se divide el número de no conformidades de cada unidad de inspección por  $n$  (número de unidades de producción en cada unidad de inspección). En el ejemplo  $n = 10$ .

Con esta información se calcula el promedio y la desviación estándar del número de no conformidades por unidad como;  $\bar{u} = \sum_{i=1}^N u_i / N$  y  $s = \sqrt{\bar{u}/n}$ .

Ahora es posible calcular los límites de control para el gráfico  $u$ , de la siguiente manera:

$$LIC = \bar{u} - 3\sqrt{\bar{u}/n} \quad ; \quad LC = \bar{u} \quad ; \quad LSC = \bar{u} + 3\sqrt{\bar{u}/n} \quad 2.8$$

donde  $\bar{u}$  representa el número promedio de no conformidades por unidad de inspección en el conjunto de datos preliminares.

## 2.6 Eficacia estadística de los gráficos de control

El objetivo básico de un gráfico de control es detectar, de la forma más rápida posible, cambios en el proceso. Cuando un punto cae fuera de los límites de control se dice que el proceso está fuera de control, que existe una causa asignable. Realmente, como en cualquier contraste de hipótesis estadística, existe una probabilidad de establecer una situación de fuera de control cuando el proceso realmente está bajo control (error tipo I) así como una probabilidad de decir que el proceso está bajo control (puntos entre límites de control) cuando realmente el proceso está fuera de control (error tipo II). Al diseñar un gráfico se deben tener en cuenta estos dos tipos de errores que determinan la eficacia estadística del mismo. Dicha eficacia se mide mediante dos indicadores; la longitud promedio de corrida o curva ARL y la función característica de operación o curva OC.

Básicamente, el ARL (Average Run Length), es el número promedio de puntos que deben graficarse antes de que un punto indique una condición de fuera de control. Es la principal característica del gráfico ya que mide la rapidez de respuesta del mismo.

Si las observaciones del proceso no están correlacionada entonces, para cualquier gráfico de control de Shewhart, el ARL se obtiene de la siguiente manera:

$$ARL = 1/p \quad 2.9$$

donde  $p$  es la probabilidad de que cualquiera de los puntos exceda los límites de control.

Cuando el proceso se encuentra bajo control, la longitud promedio de corrida se conoce como  $ARL_0$ , mientras que si el proceso se encuentra fuera de control, se conoce como  $ARL_1$ . En particular,  $ARL_0 = 1/\alpha$  y  $ARL_1 = 1/1 - \beta$ .

Por ejemplo, en el caso de los gráficos de control para la fracción de no conformes, las probabilidades  $\alpha$  y  $\beta$  pueden calcularse directamente a partir de la distribución binomial del siguiente modo:

$\alpha =$  probabilidad de rechazar un lote cuando  $p = p_0$

$$\alpha = Pr(\hat{p} > LSC/p = p_0) + Pr(\hat{p} \leq LIC/p = p_0) \quad 2.10$$

$$= 1 - \{Pr(\hat{p} \leq LSC/p = p_0) - Pr(\hat{p} \leq LIC/p = p_0)\}$$

$\beta =$  probabilidad de no rechazar un lote cuando  $p = p_1$

$$\beta = Pr(\hat{p} \leq LSC/p = p_1) - Pr(\hat{p} \leq LIC/p = p_1) \quad 2.11$$

donde  $p_0$  es la proporción de unidades no conformes en la muestra cuando el proceso está bajo control y  $p_1$  es la proporción de unidades no conformes en la muestra cuando el proceso está fuera de control.

De igual manera pueden obtenerse los valores de  $\alpha$  y  $\beta$  para los otros gráficos de control.

En general, para cualquier gráfico de control con límites  $3\sigma$  ( $\alpha = 0,0027$ ) se cumple que, si el proceso se encuentra bajo control, se experimentará una señal de falsa alarma aproximadamente cada 370 muestras ( $ARL_0 = 1/0,0027 = 370$ ).

Cuando el proceso se encuentra fuera de control, se debe prestar atención al valor de  $\beta$ . En general, cuanto más alejado se encuentre  $p_1$  de  $p_0$  mayor será la potencia de la prueba ( $1 - \beta$ ) y, en consecuencia,  $ARL_1$  será menor.

Otra medida de sensibilidad del gráfico de control es la curva OC. La curva OC, también conocida como función de operación característica, es una representación gráfica de la probabilidad de aceptar incorrectamente la hipótesis de que el proceso se encuentra bajo control estadístico (error tipo II).

Luego, dados un par de límites de control y un valor de la estadística de interés, la curva OC permite detectar el corrimiento de la estadística respecto a su valor nominal.

## 2.7 Algunos problemas relacionados con los gráficos de atributos

El control estadístico de procesos, basado en los gráficos de Shewhart, ha sido ampliamente usado en la industria manufacturera debido a la simplicidad de su implementación y a la facilidad en la interpretación del estado del proceso. No obstante, el acelerado desarrollo de la tecnología, ha motivado la mejora de los procesos hasta un punto tal que muchos gráficos de control presentan actualmente problemas de implementación práctica. Esta situación es más seria en el caso de los gráficos de atributos, cada día más utilizados, debido a la posibilidad de obtener rápidamente datos de conteos permitiendo a los procesos ser monitoreados a bajo costo. Varios autores, entre los que se destacan, Montgomery (2008), Acosta-Mejía (1999), Xie *et al.* (2002), entre muchos otros, han considerado algunos problemas asociados con el uso de los gráficos de atributos cuando los procesos tienden a un entorno cercano a cero no conformidades. Entre las principales cuestiones a tener en cuenta, destacan las siguientes;

- Por una parte, los tradicionales gráficos  $p$  y  $np$  están basados en la aproximación normal. Cuando  $p$  es pequeño y  $n$  no es suficientemente grande, la probabilidad de falsa alarma se torna muy elevada. En esta situación, los límites de control basados en la aproximación normal, pueden no describir adecuadamente la situación real del estado del proceso ni detectar correctamente los cambios. Un problema relacionado con esta situación es que debido a la alta probabilidad de falsas alarmas, las verdaderas señales pueden ser pasadas por alto.
- Otro problema se presenta en las cercanías de cero defectos o cero no conformidades. Cuando el proceso opera con alta calidad, es decir, cuando muchas muestras carecen de defectos, suele suceder que el límite inferior de control es menor a cero lo que dificulta seriamente la interpretación del comportamiento del proceso. Las mejoras no pueden ser detectadas fácilmente y la lectura de las corridas son complejas de entender.
- No menos serio es el problema que se origina con el límite superior de control. Cuando el proceso opera con muy bajo número de defectos, el límite superior de control puede ser menor que uno, salvo que el tamaño de la muestra sea muy grande. En estos casos, cualquier muestra que contenga sólo un artículo no conforme causará una señal de falsa alarma, invalidando el control en sí mismo.

- Una situación similar sucede con los gráficos  $c$  y  $u$ . Si  $c$  es chico, el límite superior de control también puede ser menor a uno. Sin embargo, el mínimo número de no conformidades para una señal de alarma es uno, y por lo tanto, los límites del gráfico de control serán poco realistas. Inclusive, a causa de la falta de normalidad, se puede obtener una probabilidad de falsa alarma mayor a la deseada para un amplio rango de valores de  $c$ .

## **2.8 Necesidad de buscar nuevos procedimientos para el control de atributos**

Actualmente los procedimientos de control deben adaptarse a procesos con bajo número de defectos. En estas circunstancias los tradicionales gráficos de Shewhart basados en la aproximación normal, han dejado de ser apropiados, debido a que incrementan de manera significativa el riesgo de falsas alarmas. Esta situación motiva la necesidad de desarrollar una serie de alternativas más potentes que las clásicas para el control de procesos por atributos.

Algunos de estos nuevos procedimientos serán tratados en los Capítulos siguientes.

## Capítulo 3

### Métodos alternativos para gráficos de control en ausencia de normalidad

#### 3.1 Introducción

Cuando Shewhart (1926) desarrolló los gráficos de control, probablemente nunca pensó que el número de unidades no conformes o que el promedio de no conformidades por unidad de producción, podrían tomar valores muy pequeños. Actualmente, la situación de bajos niveles de no conformidades existe y los clásicos gráficos de control de atributos se tornan inadecuados debido a que incrementan de manera significativa el riesgo de falsas alarmas.

Cuando la fracción de no conformidades es reducida, la aproximación normal deja de ser adecuada y deberían utilizarse límites de probabilidad de manera de reducir la probabilidad de falsas alarmas (Xie y Goh, 1993).

Sin embargo, la determinación de límites probabilísticos en los gráficos de control, depende en gran medida del conocimiento de la distribución subyacente de la característica de calidad o de su adecuada estimación.

Cuando se usan los límites de control  $3\sigma$  tradicionales, la probabilidad de falsa alarma es de aproximadamente un 0,0027, por lo que, cuando el proceso se encuentra bajo control, los límites superior e inferior del gráfico de control deben ser obtenidos de tal manera que se mantenga este nivel de confianza asociado con el concepto  $3\sigma$  de Shewhart.

Con la intención de buscar nuevos procedimientos para tratar con procesos de este tipo, varios autores han propuesto algunos métodos alternativos.

Inicialmente fueron tres los procedimientos propuestos: a) el uso de límites de control exactos de manera que la probabilidad de falsa alarma se encuentre tan próxima como sea posible al nivel  $\alpha$  prefijado; b) la utilización de transformaciones que normalicen los datos y c) la realización de ajustes sobre los límites de control.

Veremos algunos de estos procedimientos que han permitido mejorar la interpretación de las señales de falsa alarma en los gráficos de control de atributos.

### 3.2 Límites de probabilidad

Se conoce que el número de unidades conformes en la muestra o alternativamente el número de unidades no conformes puede ser modelado mediante una distribución binomial o Poisson. Estas distribuciones suelen ser aproximadas por la distribución normal cuando se usan los límites  $3\sigma$ . No obstante, estos límites no deberían ser utilizados mecánicamente sin una clara justificación, especialmente cuando  $p$  o  $\lambda$  toman valores cercanos a 0 y el tamaño de la muestra es chico.

Para derivar los límites de probabilidad, supóngase que  $Z$  es la medición que queremos graficar, ya sea que corresponde a un promedio o a un conteo de unidades no conformes. Si  $F(Z)$  es la función de probabilidad acumulada de  $Z$ , el límite de probabilidad inferior  $z_l$  y el límite de probabilidad superior  $z_s$  se obtienen como solución o aproximación de las siguientes ecuaciones:

$$\begin{aligned} P(Z \leq z_l) &= F(z_l) = \alpha/2 \\ P(Z \leq z_s) &= F(z_s) = 1 - \alpha/2 \end{aligned} \tag{3.1}$$

donde  $\alpha$  es el nivel prefijado de la probabilidad de falsa alarma.

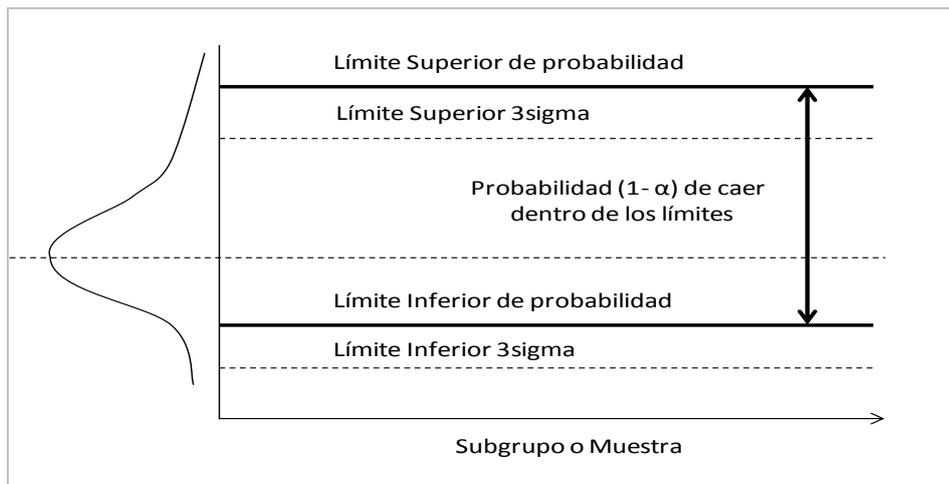


Figura 3.1: Límites de probabilidad usados cuando la característica no es normal.

Cuando se usan los gráficos  $np$ , dado que la distribución subyacente es binomial, se puede calcular la probabilidad exacta de que el número de unidades no conformes  $Z$  sea menor que, o mayor que un número dado. Los límites de control deberían fijarse de acuerdo a estas probabilidades. Por simplicidad, supóngase que todas las muestras son del mismo tamaño. Luego, la probabilidad de encontrar  $k$  o menos unidades no conformes en la muestra de tamaño  $n$  se obtiene como:

$$P(Z \leq k) = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}, \quad k = 0, 1, 2, \dots, n \quad 3.2$$

Cuando esta probabilidad es igual a 0,00135 se obtiene el límite de control inferior y cuando es igual a 0,99865 el límite de control superior. Estos valores corresponden a los niveles  $3\sigma$  para el caso en que la aproximación normal sea la adecuada. Con las herramientas informáticas actuales, estos cálculos se obtienen rápidamente.

Sin embargo, se debe notar que, como la distribución binomial es discreta, no es siempre posible especificar el nivel de confianza con exactitud, dado que  $k$  sólo puede tomar valores enteros. En esta situación conviene tomar el límite inferior de control como aquel valor de  $k$  para el cual:

$$\text{LIC} = \text{máx} \left\{ k: \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} \leq 0,00135 \right\} \quad 3.3$$

Nuevamente, cuando esta probabilidad es menor o igual a 0,00135 se obtiene la probabilidad de falsa alarma asociada con el límite inferior del gráfico  $3\sigma$  de la distribución normal.

En el caso de los gráficos  $c$ , se produce una situación similar. Considerando sólo el límite superior de control, el número de no conformidades necesarias para considerar el estado de fuera de control del proceso al nivel  $\alpha$  se puede calcular de la siguiente manera:

$$P(\text{de } k \text{ o más no conformidades en la muestra}) \leq \alpha \quad 3.4$$

Como el número de no conformidades en la muestra sigue una distribución Poisson con parámetro  $c$ , el límite superior de control para el gráfico se calcula con el menor valor de  $k$  que satisface la siguiente ecuación:

$$P(Z < k) = \sum_{i=0}^{k-1} \frac{e^{-c} c^i}{i!} > 1 - \alpha \quad 3.5$$

que se puede resolver fácilmente dado que  $k$  es por lo general un entero pequeño.

Aunque la ventaja de utilizar límites probabilísticos es evidente, hay algunas limitaciones a tener en cuenta. Los límites de probabilidad no constituyen una herramienta habitual en el control de procesos y requieren del conocimiento de la distribución subyacente de la característica a medir. En efecto, aunque el uso de límites de probabilidad debería ser utilizado cuando se implementa un

gráfico de control, especialmente para cantidades que no siguen una distribución normal, su utilidad para los gráficos  $np$  y  $c$ , es limitada.

Para los gráficos  $np$  y  $c$ , cuando el proceso ha mejorado, es común que el límite inferior de control sea menor a cero a causa de que el tamaño de la muestra, por lo general, por cuestiones prácticas, también es pequeño. En esta situación, no es posible detectar las mejoras del proceso a menos de que se implementen ciertas reglas, lo que hace más compleja la interpretación del gráfico. Cuando el proceso es de muy alta calidad, el límite superior de control también puede ser menor a uno, como es el caso de los límites  $3\sigma$ . En esta situación, cualquier unidad no conforme que aparezca en el proceso, llevará a la conclusión de que el proceso se encuentra fuera de control. Esto hace que los gráficos de control pierdan su utilidad para monitorear el proceso.

Otro problema práctico es que los límites probabilísticos generalmente no son simétricos lo que dificulta enormemente la interpretación del gráfico. El hecho de utilizar límites simétricos para distribuciones asimétricas, también puede ser engañoso.

En síntesis, cuando la característica del proceso no sigue una distribución normal, deberían utilizarse límites probabilísticos en lugar de los tradicionales límites  $3\sigma$ . Los límites de probabilidad se pueden obtener fijando la probabilidad de falsa alarma en un nivel aceptable y conociendo la distribución de probabilidad asociada con la característica de calidad. Sólo cuando es posible comprobar la normalidad de los datos es adecuado emplear los gráficos de control de Shewhart.

Algunas veces conviene transformar los datos de manera de normalizarlos. A continuación se muestran algunos de estos procedimientos.

### **3.3 Transformación de los datos**

El enfoque basado en la transformación de datos es otra opción para mejorar el comportamiento de los gráficos de control de atributos. Básicamente consiste en transformar una distribución asimétrica a una casi simétrica. Hay muchos enfoques estadísticos que han sido aplicados tanto a la distribución binomial como a la distribución de Poisson de manera de obtener la transformación que mejor aproxime los datos a la distribución normal.

Sin embargo, aunque la transformación de datos puede proporcionar buenos resultados, hay algunos puntos que deben considerarse. Por ejemplo, el enfoque transformador de datos requiere inspectores entrenados y tiempo para transformar los datos. Además, los gráficos no ilustran el nivel exacto de defectos en un proceso, ya que la escala original de los datos no es la real y esto dificulta la interpretación de las señales de inestabilidad.

Xie y Goh, (1993), Woodall (1997), Ryan (1989) y Quesenberry (1991), entre otros, estudiaron las propiedades de los gráficos  $p$  cuando  $p$  es conocido y propusieron gráficos alternativos basados en transformaciones. Ellos determinaron que la probabilidad de falsa alarma de los gráficos  $p$  depende fuertemente de los valores de  $n$  y de  $p$ , y que, cuando uno o ambos son muy chicos, estas probabilidades difieren considerablemente del valor nominal de  $\alpha$  (0,0027) en la distribución normal, para límites  $3\sigma$ .

Chen y Cheng (1998), presentaron algunos gráficos alternativos considerados en la literatura para tratar con el monitoreo de los gráficos  $c$  y  $u$ . En particular, es conocido que el gráfico  $u$  para monitorear  $\bar{u} = \lambda$ , el número promedio de no conformidades por unidad de inspección, requiere que  $\lambda$  sea grande ( $\lambda \geq 15$ ) de manera que el límite inferior de control sea positivo y que la probabilidad de falsa alarma no difiera considerablemente del valor de la normal.

Surath Aebtarm y Nizar Bouguila (2011) realizaron una recopilación de diferentes gráficos alternativos para el monitoreo de no conformidades en gráficos de control de atributos principalmente enfocado a los gráficos  $c$  de Poisson.

Algunos de los procedimientos propuestos son:

### 3.3.1 Gráfico Q binomial

Para  $i = 1, 2, \dots$  sea  $x_i$  una observación de una variable aleatoria binomial con parámetros  $n$  y  $p$ . El gráfico  $Q$  (Quesenberry, 1991, 1995) se basa en la transformación de esta variable binomial a una variable normal estándar. La transformación es:

$$u_i = \sum_{k=0}^{x_i} \binom{n}{k} p^k (1-p)^{n-k} \quad \text{con} \quad Q_i = \Phi^{-1}(u_i) \quad i = 1, 2, \dots \quad 3.6$$

donde  $\Phi^{-1}(\cdot)$  es la función inversa de la distribución normal estándar.

Luego, se pueden graficar los valores  $Q_1, Q_2, \dots$  sobre un gráfico con línea central dada por;  $LC = 0$  y límites de control:  $LSC = 3$  y  $LIC = -3$ . Transformando los valores de  $x_i$  mediante la función binomial, los  $u_i$  son aproximadamente uniformes entre 0 y 1, y los  $Q_i$  son aproximadamente normales.

Cuando  $p$  es desconocido, se definen  $s_i = \sum_{j=1}^i n_j$  y  $t_i = \sum_{j=1}^i x_j$  y sea:

$$u_i = \sum_{k=h_i}^{x_i} \binom{n_i}{k} \binom{t_i-k}{s_i-1} \binom{n_i+s_i-1}{t_i} \quad \text{donde} \quad h_i = \max(0, t_i - s_{i-1}) \quad \text{y} \quad Q_i = \Phi^{-1}(u_i) \quad 3.7$$

Luego, se pueden graficar los valores  $Q_1, Q_2, \dots$  sobre un gráfico con línea central dada por;  $LC = 0$  y límites de control:  $LSC = 3$  y  $LIC = -3$ . En este caso la transformación de la variable  $x_i$  se

realiza mediante la distribución hipergeométrica. Como  $u_i$  se define usando  $s_{i-1}$ , se puede comenzar a graficar sólo cuando el segundo valor está disponible.

Quesenberry (1991), estableció que, en general los gráficos  $Q$  tienen colas de probabilidad menores a 0,00135 para la cola inferior y mayores a 0,00135 para la cola superior. Concluyó además, que la transformación arcoseno del gráfico  $p$  da una mejor aproximación al área de la cola inferior nominal y que el gráfico  $Q$  determina una mejor aproximación al área de la cola superior nominal.

Básicamente, tanto el arcoseno del gráfico  $p$  como el gráfico  $Q$ , representan una estadística transformada en lugar de la estadística de interés y el procedimiento de cálculo para ambas no es sencillo.

### 3.3.2 Arcoseno del gráfico $p$

Otra transformación no lineal es la transformación arcoseno propuesta por Ryan y Schwertman (1997) que se obtiene como:

$$w_i = 2\sqrt{n_i} \left[ \text{sen}^{-1} \left( \sqrt{\frac{x_i + 3/8}{n_i + 3/4}} \right) - \text{sen}^{-1}(\sqrt{p}) \right] \quad 3.8$$

Siendo  $x_i$  el número de unidades no conformes y  $n_i$  el tamaño de la  $i$ -ésima muestra. Dado que  $w_i$  es aproximadamente normal estándar, se puede construir un gráfico de control con los valores de  $w_i$ , con límites de control superior e inferior iguales a 3 y -3 respectivamente.

Cuando  $p$  es desconocido se lo estiman tomando  $p = \hat{p}_{i-1}$  y se lo reemplaza en la ecuación 3.8.

Los autores establecieron además que el arcoseno del gráfico  $p$  tiene dos efectos beneficiosos: 1) el área de las colas estará siempre más cercana a las áreas de las colas bajo normalidad y 2) para un valor dado de  $p$ , el mínimo tamaño de muestra necesario para obtener un límite de control inferior (LIC) positivo es mucho menor que el que se obtiene con el clásico gráfico  $p$  para límites  $3\sigma$ .

Estos dos gráficos usan transformaciones no lineales para una mejor aproximación a la distribución normal.

### 3.3.3 Bartlett, Anscombe, Freeman y Tukey

Para gráficos  $c$ , Bartlett (1936), Anscombe (1948) y Freeman y Tukey (1950), propusieron tres enfoques para transformación de datos que se muestran en la Tabla 3.1. En esta tabla,  $c$  es el número de defectos (datos originales) e  $y$  es el dato transformado con varianza unitaria. Mediante estas tres transformaciones, los límites de control vienen dados por lo siguiente ecuación:

$$\text{LIC} = \bar{y} - 3 \qquad \text{LC} = \bar{y} \qquad \text{LSC} = \bar{y} + 3 \qquad 3.9$$

donde  $\bar{y}$  es la media de los datos transformados.

Tabla 3.1: Resumen de ecuaciones de transformación

Autor	Ecuación de transformación
Bartlett (1946)	$y = 2\sqrt{c}$
Ascombe (1948)	$y = 2\sqrt{c + 3/8}$
Freeman y Tukey (1950)	$y = \sqrt{c} + \sqrt{c + 1}$

### 3.3.4 Gráfico $Q$ Poisson

Como antes, sea  $\Phi^{-1}(\cdot)$  la función inversa de una distribución normal estándar, y sea  $[x]$  el mayor entero menor que  $x$ . Si se define:

$$F(x; \lambda) = \sum_{k=0}^{[x]} e^{-\lambda} \lambda^k / k! \quad , \quad x \geq 0 \qquad 3.10$$

Para  $i = 1, 2, \dots$  sea  $y_i$  una observación de una variable aleatoria  $Y_i$  con distribución Poisson, con media  $n_i \lambda$ . Quesenberry (1991), introduce los siguientes gráficos  $Q$  para el caso  $\lambda$  ( $c$ ) conocido y desconocido.

En el caso  $\lambda$  conocido, se define  $v_i = F(y_i; n_i \lambda)$  y  $Q_i = \Phi^{-1}(v_i)$ . Luego, se pueden graficar los valores  $Q_1, Q_2, \dots$  sobre un gráfico con línea central  $\text{LC} = 0$  y límites de control dados por  $\text{LSC} = 3$  y  $\text{LIC} = -3$ .

Para el caso  $\lambda$  desconocido, se define  $t_i = \sum_{j=1}^i y_j$  y  $N_i = \sum_{j=1}^i n_j$ . Para  $i = 1, 2, \dots$  y sea;  $v_i = B(y_i; t_i, n_i / N_i)$  una variable binomial, y  $Q_i = \Phi^{-1}(v_i)$ . Luego, se grafican los valores  $Q_1, Q_2, \dots$  sobre un gráfico con línea central  $\text{LC} = 0$  y límites de control dados por  $\text{LSC} = 3$  y  $\text{LIC} = -3$ .

### 3.3.5 Transformación raíz cuadrada

Otra transformación es la realizada con la función raíz cuadrada del gráfico  $c$  cuando  $\lambda$  ( $c$ ) es conocido y cuando es desconocido.

La típica transformación para estabilizar la varianza para variables aleatorias con distribución Poisson es  $2(Y_i + 3/8)^{1/2}$ , propuesta por Anscombe (1948) que tiene distribución aproximadamente normal con media  $2(n_i \lambda_i + 3/8)^{1/2}$  y varianza unitaria.

Luego,

$$W = 2\sqrt{y_i + 3/8} - 2\sqrt{n_i\lambda + 3/8} \quad 3.11$$

sustituyendo  $\lambda$  por  $\hat{\lambda}_{i-1}$  en el caso  $\lambda$  no conocida.

Quesenberry (1991) utilizó una propuesta más general, dada por;

$$W = 2\sqrt{c} - 2\sqrt{n\bar{c}} \quad 3.12$$

Con esta transformación, los datos se ubican en un gráfico de control con límites  $\pm 3$  y línea central cero.

### 3.3.6 Transformación raíz cuadrada mejorada (ISRT)

Un gráfico de control de atributo alternativo basado en la transformación mejorada de la raíz cuadrada (ISRT) fue propuesto en Tsai *et al.* (2006). Los autores utilizaron la raíz cuadrada de los datos para transformar distribuciones asimétricas, como la binomial y la Poisson, en distribuciones simétricas.

En el caso de los gráficos  $c$ , los límites de control se obtienen de la siguiente manera:

$$\text{LIC} = \sqrt{c} - \frac{3}{2} - \frac{9}{8} \left( \frac{1}{\sqrt{c}} \right) \quad \text{LC} = \sqrt{c} \quad \text{LSC} = \sqrt{c} + \frac{3}{2} - \frac{1}{2} \left( \frac{1}{\sqrt{c}} \right) \quad 3.13$$

Estas son algunas de las transformaciones que más se han utilizado para normalizar datos y facilitar la interpretación de las señales de inestabilidad de los gráficos de control.

Sin embargo, los procedimientos de transformación no fueron muy bien recibidos por los responsables de calidad debido principalmente a la dificultad relacionada con la transformación en sí misma y con la interpretación de las señales de inestabilidad y del seguimiento del proceso.

Por este motivo, algunos investigadores desarrollaron nuevos procedimientos de monitoreo y control de procesos que se encuentran alejados de la distribución normal, buscando la solución en la modificación, no ya de los datos, sino en los límites de control.

## 3.4 Transformación sobre los límites de control de los gráfico $p$ y $np$

Como se ha establecido una forma de mejorar los gráficos de atributos es mediante alguna transformación efectuada a los datos que logre una mayor potencia para detectar desplazamientos del proceso. Sin embargo, diferentes autores entre los que se encuentran, Winterbottom (1993), Xie *et al.* (2002), Xie y Goh (1995), Kuralmani *et al.* (2002), Chen y Cheng (1998), Cavalcanti y

Cordeiro (2006), entre otros, han demostrado que con un simple ajuste efectuado en los límites de control del gráfico, se pueden lograr iguales o aún mejores resultados sin modificar la escala original de las mediciones.

Una aproximación sistemática para mejorar la exactitud probabilística de los límites de control para todos los gráficos de atributos consiste en determinar el ajuste utilizando la expansión de Cornish-Fisher. Este ajuste o corrección depende de manera simple, del tamaño de la muestra, del valor de los parámetros del proceso y del valor de la amplitud  $k$  utilizada como multiplicador de  $\sigma$  en la distribución normal estándar.

La expansión de Cornish-Fisher es una fórmula para aproximar cuantiles de una variable aleatoria basada en sus cumulantes. Los cumulantes de una variable aleatoria son similares a sus momentos y se definen como aquellos valores de  $K_r$  tales que la identidad;  $\exp(\sum_{r=1}^{\infty} K_r t^r / r!) = \sum_{r=0}^{\infty} (E(X^r) t^r) / r!$ , se mantiene para todo  $t$ .

Los cumulantes de una variable aleatoria  $X$  pueden ser expresados en términos de su media,  $\mu = E(X)$  y momentos centrales  $\mu_r = E[(X - \mu)^r]$ . La expresión para los primeros cinco cumulantes es:  $K_1 = \mu$ ,  $K_2 = \mu_2$ ,  $K_3 = \mu_3$ ,  $K_4 = \mu_4 - 3\mu_2^2$  y  $K_5 = \mu_5 - 10\mu_3\mu_2$ . (Anexo I).

Cornish y Fisher (1937) desarrollaron una expansión para aproximar el cuantil  $q$ ,  $\Phi_X^{-1}(q)$  de una variable  $X$  normal con media 0 y desviación estándar unitaria, basada en sus primeros cumulantes.

Para aplicar la expansión de Cornish-Fisher a los gráficos  $p$  y  $u$ , es necesario efectuar previamente el cálculo de los momentos y cumulantes de la distribución respectiva. Para el caso de los gráficos  $p$ , este cálculo se encuentra en el Anexo I.

En las secciones siguientes, veremos la aplicación de la expansión de Cornish-Fisher para la obtención de límites de control modificados en los gráficos de control de atributos. Primero se considera la aplicación a los gráficos  $p$  y  $np$  y luego a los gráficos  $u$ .

### 3.4.1 Límites de control con un ajuste en la expansión de Cornish-Fisher

Winterbottom (1993) logró una significativa mejora en los límites de control mediante un simple ajuste (cumulante de 3° orden) derivado de la expansión cuantílica de Cornish-Fisher. Determinó que este ajuste era más adecuado para normalizar los datos que las transformaciones, debido a que se retiene la escala original de las mediciones. Winterbottom llevó a cabo su estudio para gráficos de atributos mediante la incorporación de un término de ajuste en la expansión asintótica de Cornish-Fisher para una corrección de sesgo y asimetría. Posteriormente Chen (1998), extendió los resultados de Winterbottom para  $p$  conocido al caso  $p$  desconocido.

El procedimiento empleado por Winterbottom para obtener la modificación de los límites del gráfico de control, se basaron en el desarrollo siguiente.

Si  $X$  es una variable aleatoria con distribución binomial de parámetros  $n$  y  $p$ , siendo  $n$  el tamaño de la muestra y  $p$  la proporción de unidades defectuosos o no-conformes en un proceso, entonces,  $Y = X/n$  es una proporción binomial con;  $\mu = E(X) = p$ , y momentos centrales;  $\mu_2 = V(X) = p(1-p)/n$  y  $\mu_3 = p(1-p)(1-2p)/n^2$  (Ecuación I.4, del Anexo I).

Sea  $z_\alpha$  el  $\alpha$ -ésimo cuantil de la distribución normal estándar, luego el  $\alpha$ -ésimo cuantil de  $Y$ , determinado  $Y_\alpha$  (que se llamará aquí  $Y_\alpha(1)$ ), obtenido de la expansión de Cornish-Fisher (Cornish & Fisher, 1960; Lee & Lee, 1992) está dado por:

$$\frac{Y_\alpha(1) - p}{\sqrt{\frac{p(1-p)}{n}}} \cong z_\alpha + \frac{(z_\alpha^2 - 1)}{6} K_3 \quad 3.14$$

donde  $K_3 = \mu_3^* = \mu_3/\sigma_3$ , (Ecuación I.7)

Se sigue inmediatamente que para  $p$ :

$$Y_\alpha(1) = p + z_\alpha [p(1-p)/n]^{1/2} + \frac{(z_\alpha^2 - 1)}{6n} (1 - 2p) \quad 3.15$$

y para  $np$ :

$$Y_\alpha(1) = np + z_\alpha [np(1-p)]^{1/2} + \frac{(z_\alpha^2 - 1)}{6} (1 - 2p) \quad 3.16$$

Fijando  $z_\alpha = \pm 3$  se obtienen los límites de control  $3\sigma$  ajustados para el gráfico  $p$  de la siguiente manera:

$$LSC_1 = p + 3 [p(1-p)/n]^{1/2} + \frac{4}{3n} (1 - 2p) = LSC + \frac{4}{3n} (1 - 2p) \quad 3.17$$

y

$$LIC_1 = p - 3 [p(1-p)/n]^{1/2} + \frac{4}{3n} (1 - 2p) = LIC + \frac{4}{3n} (1 - 2p) \quad 3.18$$

Se ha llamado a estos límites  $LSC_1$  y  $LIC_1$  para distinguirlos de los límites de control del gráfico tradicional para  $p$ .

Nótese que para el gráfico  $p$  el término de corrección es igual a  $4(1-2p)/3n$  y para el gráfico  $np$ , es  $4(1-2p)/3$ .

Estos términos de corrección pueden ser determinados también para otros valores de  $z_\alpha$ . Por ejemplo, es posible establecer límites de control de alertas para  $z_\alpha = \pm 2$ . La corrección para límites  $2\sigma$  en el gráfico  $p$  es  $(1 - 2p)/2n$ .

A continuación se presenta un estudio comparativo de las ventajas del gráfico  $p$  modificado, propuesto por Winterbottom respecto del gráfico tradicional basado en la distribución normal, evaluados mediante el riesgo de falsa alarma  $\alpha$ .

**(i) Estudio comparativo de falsa alarma:** El gráfico  $p$  teórico ideal es aquel que tiene un riesgo  $\alpha$  prefijado para el error de Tipo I (rechazar la  $H_0$  cuando la  $H_0$  es verdadera) que es considerado usualmente, para los gráficos de control, igual a 0,0027 y que corresponde a la amplitud  $3\sigma$  de la distribución normal.

Esto indicaría que bajo distribución normal, el riesgo  $\alpha$  debería permanecer constante oscilando alrededor de 0,0027 para cualquier tamaño de muestra y fracción de no conformes. Sin embargo, es conocido en la literatura que esta situación se observa sólo cuando  $np(1 - p) > 5$  y  $0.1 < p < 0.9$ , o cuando  $np(1 - p) > 25$  (Xie et al., 2002). Por lo tanto, para cualquier combinación de  $n$  y  $p$  que no cumpla con esta regla, el valor de  $\alpha$  se alejará del valor teórico indicando que el riesgo de falsas alarmas se habrá incrementado.

Cabe aclarar que dado la naturaleza discreta de los datos, el valor de  $\alpha$  no es continuo sino que se expresa de a saltos para cada valor de  $p$  (dado  $n$ ). Además, la línea de corte ascendente que se observa para valores de  $p$  muy pequeños, indica que el riesgo de falsa alarma se incrementa de tal manera que el gráfico de control ya no es adecuado.

La idea entonces es mostrar gráficamente los beneficios de la modificación realizada sobre los límites de control, efectuando la comparación del valor que toma el riesgo  $\alpha$  para el caso normal y cuando se utilizan límites de control con un ajuste. Para ello se muestran los valores de  $\alpha$ , para el caso de límites de control bilaterales y unilaterales, considerando un tamaño de muestra fijo y variando los valores de  $p$  de manera de lograr establecer cuál es el menor valor de la fracción de no conformes para la cual el riesgo  $\alpha$  se aproxima más al valor teórico.

A modo de ejemplo, en la Figura 3.2 se muestra el riesgo de falsa alarma para el caso  $n = 20$ .

[Las referencias para ambos gráficos son: • aproximación Normal; ▲ Cornish-Fisher (un ajuste);  $\alpha_0 = 0,0027$ ]

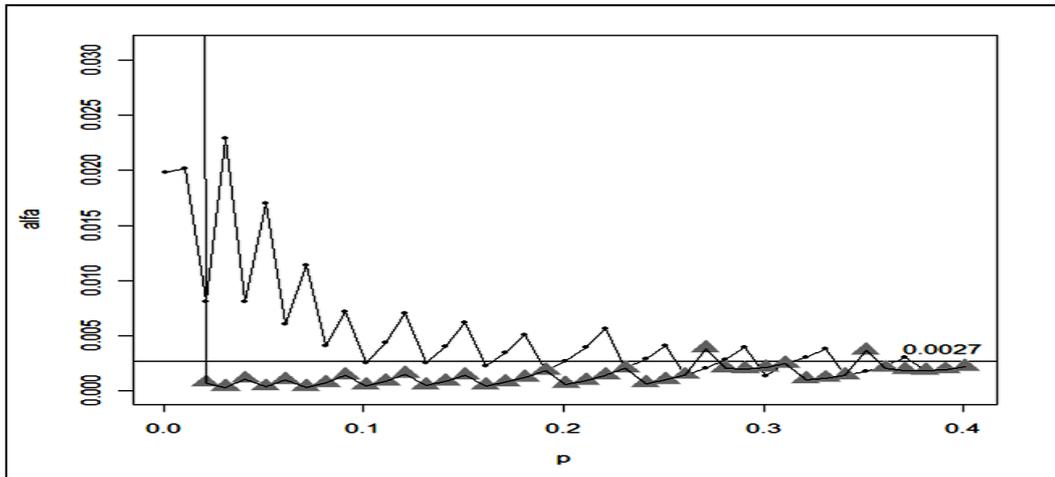


Figura 3.2: Riesgo  $\alpha$  bilateral para aproximación normal contra Cornish-Fisher (con un ajuste)

En la Figura anterior, resulta claro que el gráfico corregido de Winterbottom presenta un riesgo de falsa alarma mucho más cercano al del riesgo de referencia ( $\alpha_0 = 0,0027$ ) que el gráfico tradicional basado en la distribución normal. Sin embargo, esta corrección no se puede utilizar para cualquier valor de  $p$ , sólo para valores de  $p$  superiores a 0,02 (aproximadamente) cuando  $n = 20$ .

Si se considera sólo el riesgo superior (probabilidad de superar el límite superior de control cuando  $H_0$  es verdadera), los resultados comparativos son incluso más evidentes a favor del gráfico corregido, como se muestra a continuación en la Figura 3.3.

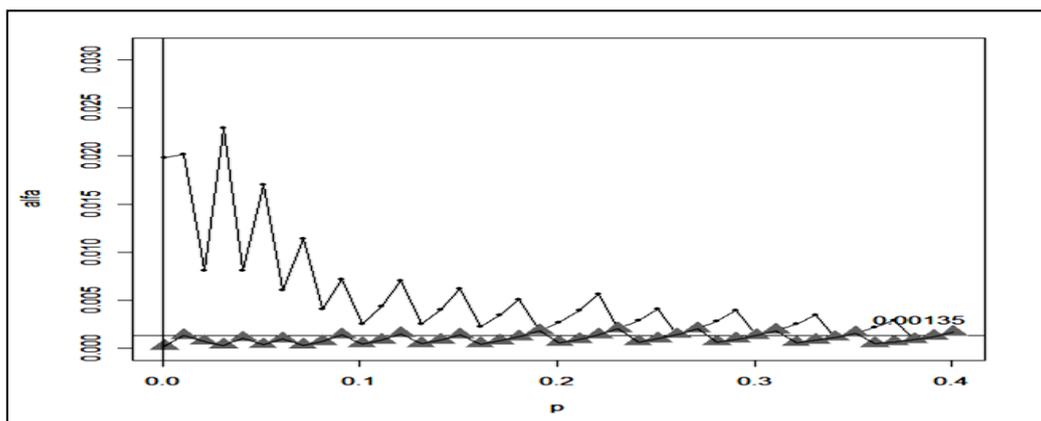


Figura 3.3: Riesgo  $\alpha$  unilateral para aproximación normal contra Cornish-Fisher (con un ajuste)

En las Figuras anteriores se puede ver que la corrección produce excelentes resultados. El exceso de falsa alarma se reduce sustancialmente respecto al caso normal.

### 3.4.2 Límites de control con dos ajustes en la expansión de Cornish-Fisher

A partir de los resultados obtenidos por la incorporación de un término de ajuste en la expansión de Cornish-Fisher, se consideró la posibilidad de incorporar un nuevo término que permitiera la corrección no sólo del sesgo y la asimetría de la distribución sino también de la curtosis. Para ello, en esta investigación, se efectuó una extensión de los resultados obtenidos por Winterbottom (1993) incluyendo términos del orden  $n^{-3/2}$  en un gráfico  $p$  mejorado.

Se pudo observar que con estos nuevos límites de control se logra una mejora importante que equivale a utilizar la distribución binomial exacta para el diseño del gráfico. Además, este gráfico supera al gráfico  $p$  propuesto por Winterbottom (que tiene una corrección del orden de  $n^{-1}$ ), en el sentido de que permite el monitoreo de procesos con valores de  $p$  más pequeños.

Para la obtención de los límites modificados se partió de la expansión cuantílica de Cornish-Fisher utilizando la información de los cumulantes de 3° y 4° orden ( $K_3$  and  $K_4$ , Ecuación I.7).

En este caso, se adicionó a la información dada en la sección anterior el momento central de 4° orden de  $Y$ ,  $\mu_4 = \{np(1-p)[1-6p(1-p)] + 3n^2p^2(1-p)^2\}/n^4$  (Ecuación I.4).

Luego el  $\alpha$ -ésimo cuantil de  $Y$ , determinado por  $Y_\alpha$  (que se llamará aquí  $Y_\alpha(2)$ ), obtenido de la expansión de Cornish-Fisher con dos términos de corrección, está dado por:

$$\frac{Y_\alpha(2) - p}{\sqrt{\frac{p(1-p)}{n}}} = z_\alpha + \frac{(z_\alpha^2 - 1)}{6} K_3 + \frac{(z_\alpha^3 - 3z_\alpha)}{24} K_4 - \frac{(2z_\alpha^3 - 5z_\alpha)}{36} K_3^2 \quad 3.19$$

(donde  $K_3 = \mu_3^* = \frac{\mu_3}{\sigma_3}$ ,  $K_4 = \mu_4^* - 3\mu_2^{*2} = \frac{\mu_4}{\sigma_4} - 3\left(\frac{\mu_2}{\sigma_2}\right)^2$ , como en Ecuación I.7)

Se sigue inmediatamente que:

$$Y_\alpha(2) = Y_\alpha(1) + \frac{(z_\alpha^3 - 3z_\alpha)[1 - 6p(1-p)]}{24n^2} \frac{1}{[p(1-p)/n]^{1/2}} - \frac{(2z_\alpha^3 - 5z_\alpha)}{36n^2} \frac{(1-2p)^2}{[p(1-p)/n]^{1/2}} \quad 3.20$$

Fijando  $z_\alpha = \pm 3$  se obtienen los límites de control  $3\sigma$  para el gráfico  $p$  mejorado de la siguiente manera:

$$LSC_2 = LSC_1 - \frac{[p(1-p) + 2]}{6n^2 [p(1-p)/n]^{1/2}} \quad 3.21$$

y

$$LIC_2 = LIC_1 - \frac{[p(1-p) + 2]}{6n^2 [p(1-p)/n]^{1/2}} \quad 3.22$$

Nótese que se ha llamado a estos límites  $LSC_2$  y  $LIC_2$  para distinguirlos de los dados en la sección anterior.

(i) **Estudio comparativo de falsa alarma:** Volviendo al ejemplo de la sección anterior, en la Figura 3.4 se muestra el riesgo comparativo de falsa alarma para  $n = 20$ .

[Las referencias para ambos gráficos son: • aproximación normal; ● Cornish-Fisher;  $\alpha_0 = 0.0027$ ]

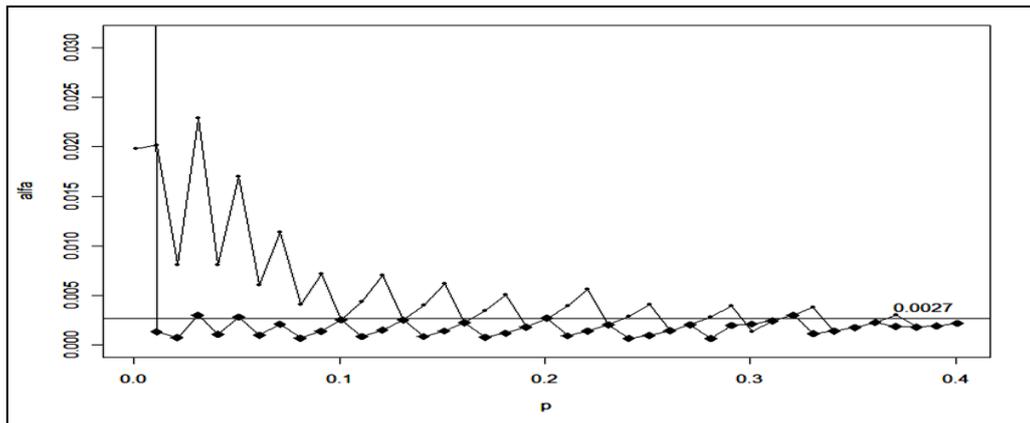


Figura 3.4: Riesgo  $\alpha$  bilateral para aproximación normal contra Cornish- Fisher (con dos ajustes)

Observando la Figura 3.4, resulta evidente que la corrección con dos términos coloca al riesgo de falsa alarma mucho más cerca del riesgo de referencia ( $\alpha_0 = 0,0027$ ) que el gráfico tradicional basado en la distribución normal. Puede notarse también que la incorporación de un término de corrección adicional, permite trabajar con valores de  $p$  menores a los obtenidos con sólo un término de corrección. Sin embargo, esta modificación en los límites de control, sólo puede utilizarse para valores de  $p$  superiores a 0,01 (aproximadamente), cuando  $n = 20$  (lo que la hace menos restrictiva que la anterior).

Si ahora se considera sólo el riesgo superior, los resultados comparativos son similares a los obtenidos para el caso bilateral, como se muestra en la Figura 3.5.

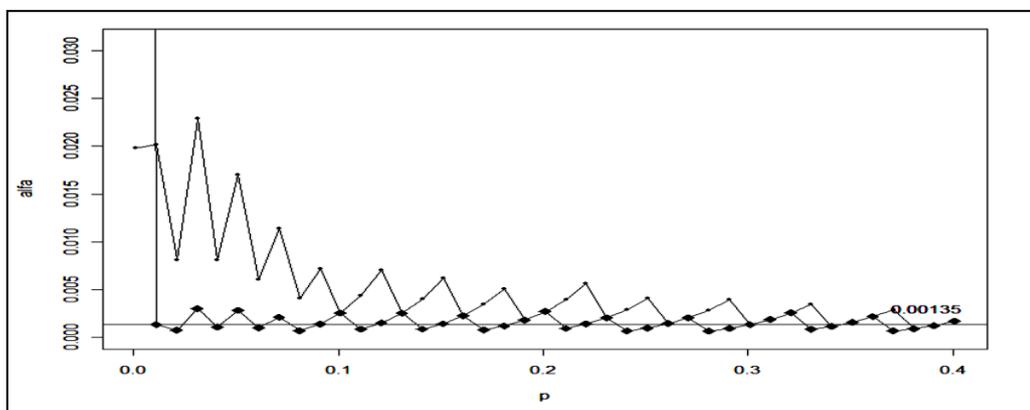


Figura 3.5: Riesgo  $\alpha$  unilateral para aproximación normal contra Cornish -Fisher (con dos ajustes)

Al comparar los resultados de ambas Figuras es posible observar que para el gráfico de control unilateral (Figura 3.5) con sólo el límite superior (situación más común en procesos con bajo número de no conformidades), el riesgo de falsa alarma se aproxima más al valor prefijado.

**(ii) Estudio de la falsa alarma para valores pequeños de  $n$  y  $p$**

A efectos de mostrar con más detalle la comparación entre los gráficos con límites modificados con uno y dos términos de corrección, en esta sección se presenta un estudio comparativo del riesgo de falsa alarma considerando tres tamaños de muestra;  $n = 20, n = 10$  y  $n = 5$ .

Para una mejor visualización de los resultados, se toman valores de  $p$  ubicados en las proximidades de 0. Esto permite observar lo que sucede con el riesgo de falsa alarma para pequeños valores de  $p$  y establecer la ventaja de incluir términos de corrección en los límites del gráfico de control.

[Las referencias para las Figuras son: ▲ Cornish-Fisher (con un ajuste), ● Cornish-Fisher (con dos ajustes);  $\alpha_0 = 0,0027$ ]

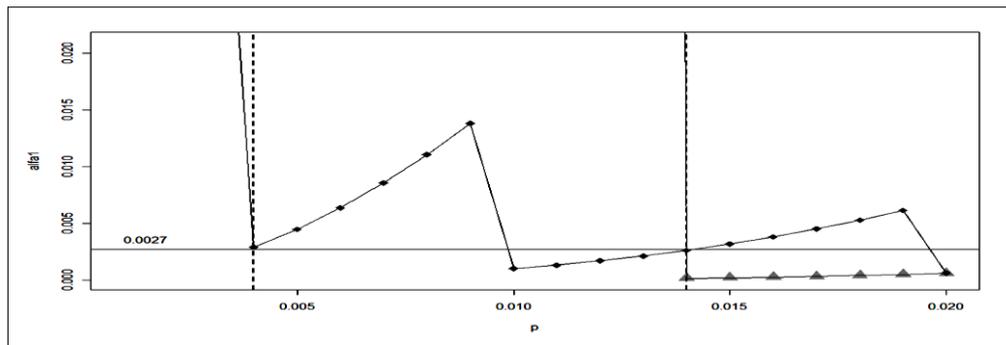


Figura 3.6: Riesgo alfa – Bilateral, ( $n = 20, p < 0,020$ )

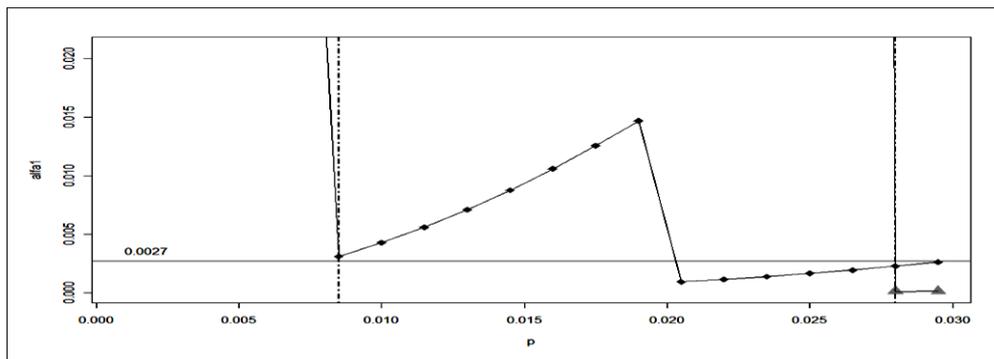


Figura 3.7: Riesgo alfa – Bilateral, ( $n = 10, p < 0,030$ )

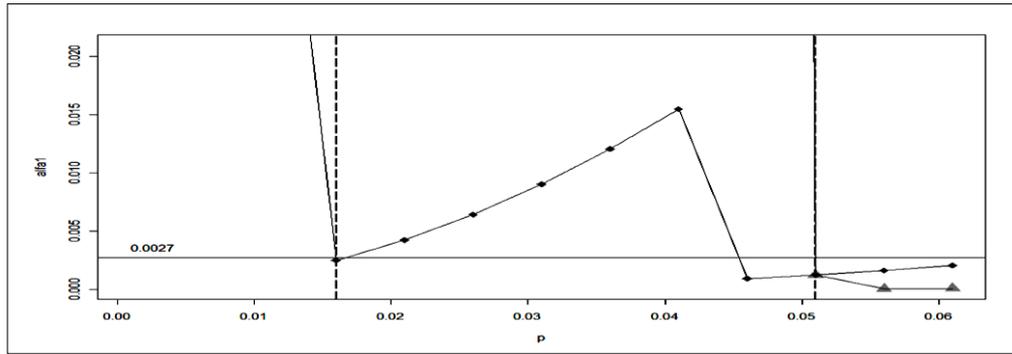


Figura 3.8: Riesgo alfa – Bilateral, ( $n = 5$ ,  $p < 0,060$ )

De las Figuras 3.6, 3.7 y 3.8, es posible observar que, aunque la modificación de la expansión de Cornish-Fisher con un término de corrección funciona bien para pequeños valores de  $p$ , la corrección con dos términos muestra mejores resultados, especialmente en las inmediaciones de cero. Nótese que la línea de puntos ascendente indica el menor valor de  $p$  para el cual el gráfico de control puede ser implementado manteniendo el riesgo de falsa alarma en las cercanías del valor teórico prefijado.

Por ejemplo, para  $n = 20$  (Figura 3.6), es posible utilizar gráficos para  $p$  con límites de control modificados con dos términos de corrección cuando  $p$  es mayor a 0,004 (0,4%), mientras que con un término de corrección sólo es posible utilizar estos gráficos cuando los valores de  $p$  superan a 0,014 (1,4%).

Para  $n = 10$  (Figura 3.7), la incorporación de un nuevo término de corrección permite trabajar con valores de  $p$  superiores a 0,008 (0,8%), mientras que la corrección con un término permite utilizar los límites modificados sólo para valores de  $p$  superiores a 0,028 (2,8%).

Cuando  $n = 5$  (Figura 3.8), la nueva modificación permite trabajar con valores de  $p$  mayores a 0,016 (1,6%), mientras que con sólo un término de corrección es posible trabajar con valores de  $p$  mayores 0,051 (5,1%).

Esta situación evidencia que, cuando el proceso opera con valores de  $p$  muy bajos (y  $n$  también es chico), una buena práctica para mantener el riesgo de falsa alarma en su valor preestablecido, consistirá en trabajar con límites de control modificados.

### (iii) Tabla de resultados

A efectos de visualizar esta información para diferentes valores de  $n$ , se construyó la Tabla 3.2. En la Tabla se muestra el mínimo valor de  $p$  que puede ser utilizado a fin de mantener el riesgo de falsa alarma bajo control, con: a) gráficos basados en la distribución normal (sin corrección), b) Cornish-Fisher con un término de corrección (CF1) y c) Cornish-Fisher con dos términos de corrección (CF2).

Tabla 3.2: Valores mínimos de  $p$  según tamaño de muestra y tipo de corrección

Tamaño muestral $n$	Minimo $p$			Minimo $np$			Minimo $np(1-p)$		
	Aprox. Normal	CF1	CF2	Aprox. Normal.	CF1	CF2	Aprox. Normal	CF1	CF2
3	-	-	0,028	-	-	0,084	-	-	0,082
5	-	0,051	0,016	-	0,255	0,080	-	0,242	0,079
10	0,410	0,028	0,008	4,100	0,280	0,080	2,419	0,272	0,079
15	0,385	0,019	0,006	5,775	0,285	0,090	3,552	0,280	0,090
20	0,360	0,014	0,004	7,200	0,280	0,080	4,608	0,276	0,080
30	0,280	0,010	0,003	8,400	0,300	0,090	6,048	0,297	0,090
40	0,240	0,008	0,002	9,600	0,320	0,080	7,296	0,317	0,080
60	0,210	0,005	0,0013	12,600	0,300	0,078	9,954	0,299	0,078
100	0,150	0,003	0,0008	15,000	0,300	0,080	12,750	0,299	0,080

La Tabla muestra los valores mínimos de  $p$  para los cuales el riesgo  $\alpha$  existe (y es finito) y está próximo al valor prefijado de  $\alpha = 0,0027$ . En general, como fue mencionado anteriormente, existe acuerdo entre diversos autores en que cuando  $np(1 - p) \geq 5$  y  $p > 0,10$ , es conveniente utilizar la aproximación normal para calcular los límites de control del gráfico  $p$  (Xie *et al.*, 2002).

Sin embargo, basado en los resultados anteriores, se puede observar que es posible extender esta regla a otros valores de  $n$  y  $p$ . De hecho, cuando  $np(1 - p) \geq 0,08$  es factible considerar la corrección CF2 para calcular los límites del gráfico de control para  $p$  y que, cuando este cálculo da un valor superior a 0,25, se puede utilizar tanto CF1 como CF2 para modificar los límites de control.

### 3.4.3 Aplicación a un proceso de manufactura

Considérese un proceso de fabricación que produce un cable del freno de mano de ciertos vehículos. El proceso debe garantizar una cierta longitud del conjunto para lo cual se ha establecido en la estación de montaje un procedimiento de control del tipo aprobado/reprobado. Cuando se determina la conformidad de la longitud del conjunto, el cable es aceptado. De lo contrario, es rechazado. La inspección se hace por muestreo tomando muestras de 20 piezas por hora de producción y contando el número de piezas no conformes. La proporción de las no conformidades del proceso ha sido monitoreada por algún tiempo empleando el gráfico  $p$  de Shewhart con  $\alpha_0 = 0,0027$  (límites  $\pm 3\sigma$ ). Actualmente el proceso está trabajando con una proporción de no conformidades de alrededor del 1,5% (15.000 ppm), y el gráfico de control ya no es eficaz porque muestra excesivas señales de falsas alarmas (incluso con el proceso bajo control).

Este ejemplo será utilizado para ilustrar el desempeño comparativo de los tres gráficos  $p$  desarrollados en este Capítulo, es decir, el gráfico tradicional de Shewhart, el gráfico con límites

corregidos con un término en la expansión de Cornish-Fisher (CF1) y el gráfico corregido con dos términos en la expansión de Cornish-Fisher (CF2).

En la Figura 3.9, se comparan las tres posibilidades metodológicas consideradas (basado en normal, CF1 y CF2). La línea vertical indica la separación entre la Fase I (100 muestras de calibración utilizadas en la estimación de  $p$ ) y la Fase II. Con el valor de  $\hat{p}$  obtenido en la Fase I se construyen los límites de control del gráfico y luego, en la Fase II, se grafican los nuevos valores de la estadística que son confrontados con los límites de control. La proporción muestral de no conformes estimada en la Fase I fue  $\hat{p} = 0,015$ .

Como referencia, LSC indica el límite superior de control del gráfico basado en la distribución normal (Shewhart), LSC<sub>1</sub> indica el límite superior de control del gráfico CF1 y LSC<sub>2</sub> indica el límite superior de control del gráfico CF2.

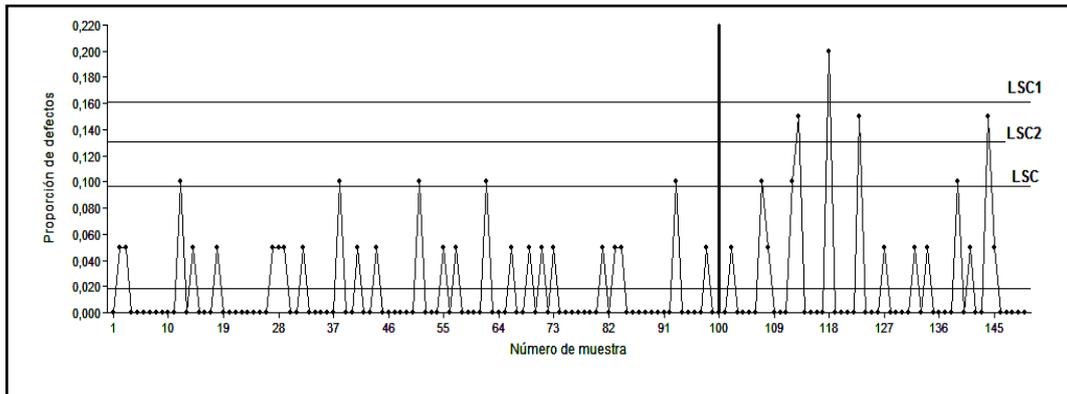


Figura 3.9: Gráfico  $p$  de control bilateral,  $n = 20$ ,  $\hat{p} = 0,015$ ,  $\alpha_0 = 0,0027$

Los límites de control y el valor del riesgo  $\alpha$  para los tres gráficos (normal, CF1 y CF2) son como se muestran en la Tabla 3.3. En este caso se ha considerado sólo el límite superior de control debido a que el límite inferior es cercano a cero o negativo.

Tabla 3.3: Límites de control y riesgo  $\alpha$  para el gráfico de la Figura 3.9

Tipo de gráfico	LSC	nLSC	Riesgo $\alpha$
Aprox. Normal.	0,0965	1,931	0,035746
CF1	0,1612	3,224	0,000202
CF2	0,1303	2,606	0,003178

Alwan (2000) señala que cuando se utiliza una distribución simétrica (normal) para aproximar una distribución binomial asimétrica (debido a los pequeños valores que toma  $p$ ), es posible que el límite inferior de control calculado tome un valor negativo. En tales casos el LIC se fija en cero, lo que implica que para todo propósito práctico, este límite no desempeña ningún rol.

En el ejemplo, debido a que el valor de  $\hat{p}$  es bastante pequeño, el límite inferior de control siempre será igual a 0. Para fines prácticos, la mejora de un proceso (valores por debajo del límite inferior de control) se determina por el valor estimado de  $p$  supervisado a lo largo del tiempo.

En la Figura 3.9, está claro que el límite más adecuado es el correspondiente a CF2 ( $LSC_2$ ) porque el basado en la normal ( $LSC$ ) muestra excesivas señales de falsa alarma y el límite basado en CF1 ( $LSC_1$ ) tiene una probabilidad de riesgo de falsa alarma prácticamente nula y muy alejada del valor prefijado en 0,0027.

Con posterioridad, la gerencia consideró necesario analizar el sistema e introducir mejoras en el proceso por considerar que la proporción unidades no conformes era inaceptablemente alta. Utilizando herramientas de calidad para resolución de problemas, (tormenta de ideas, Pareto, diagramas de causa-efecto, etc.), la empresa logró reducir la proporción de unidades con algún defecto a un valor de  $\hat{p} = 0,004$  (4000 ppm).

En la Figura 3.10, se muestra el gráfico de control para esta nueva situación.

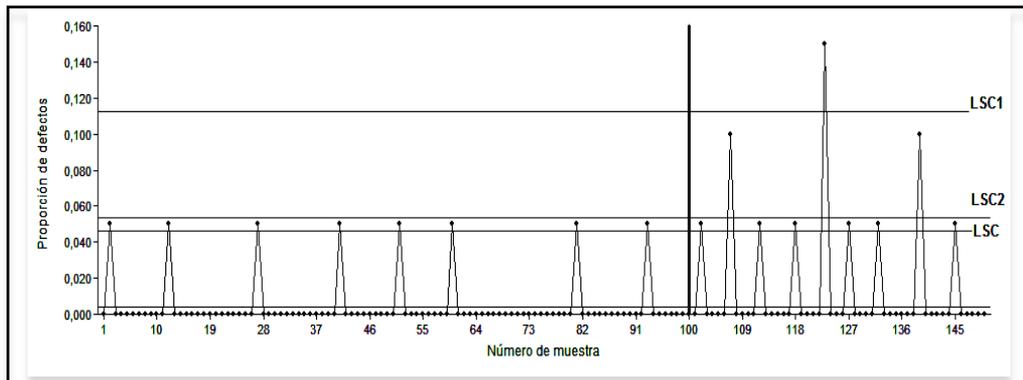


Figura 3.10: Gráfico  $p$  de control bilateral,  $n = 20$ ,  $\hat{p} = 0,004$ ,  $\alpha_0 = 0,0027$

Los límites de control y el valor del riesgo  $\alpha$  para los tres gráficos (normal, CF1 y CF2) son como se muestran en la Tabla 3.4. Nuevamente se ha considerado sólo el límite superior de control debido a que el límite inferior está muy próximo a cero o es negativo.

Tabla 3.4: Límites de Control y riesgo  $\alpha$  para el gráfico de la Figura 3.10

Tipo de gráfico	LSC	nLSC	Riesgo $\alpha$
Aprox. Normal.	0,0463	0,926	0,077032
CF1	0,1125	2,250	0,923038
CF2	0,0533	1,066	0,002898

En esta nueva situación, los límites de control basados en la normal muestran una alta probabilidad de falsas alarmas. Esto se debe principalmente al hecho de que el límite superior de

control (LSC) es menor a uno, y luego no es apropiado para controlar el proceso dado que la presencia de sólo una no conformidad indicará un estado de fuera de control. El límite CF1 tampoco debería considerarse, dado que muestra un riesgo de falsa alarma muy elevado y alejado de 0,0027. Por el contrario, el límite CF2 muestra un riesgo de falsa alarma muy cercano al prefijado, indicando que ésta es la mejor opción.

Nótese además, que en esta situación ( $n = 20$  y  $\hat{p} = 0,004$ ),  $np(1 - p)$  es igual a 0,08, que es el valor recomendado para utilizar la corrección CF2 en los límites de control según lo informado en la Tabla 3.2.

### 3.5 Transformación sobre los límites de control de los gráficos $c$ y $u$

Cuando se utilizan gráficos  $c$  relacionados con el número de defectos por unidad de inspección, la variable sigue una distribución de Poisson con parámetro  $c$  ( $\lambda$ ), como se ha presentado en el Capítulo anterior. Estos gráficos se construyen generalmente bajo el supuesto de que es adecuado aproximar la distribución normal a la distribución Poisson, suponiendo que  $c$  (o  $\lambda$ ) es conocido.

Cuando no se tiene un valor conocido para  $\lambda$ , éste puede ser estimado como el promedio de no conformidades en una muestra preliminar de unidades de inspección.

Se conoce que cuando  $\lambda \leq 20$ , el límite inferior de control (LIC) de estos gráficos, se torna poco efectivo y fracasa como indicador de una disminución del valor real  $\lambda$  del proceso.

Para dar solución a este problema, Winterbottom (1993) propuso gráficos modificados para  $c$  y  $u$ , ajustando los límites de control tradicionales con términos del orden  $n^{-1}$ , donde  $n$  es el tamaño de la muestra de inspección.

Posteriormente, Chen y Cheng (1998), retomaron el trabajo de Winterbottom efectuando una pequeña corrección teórica a sus resultados y presentando la modificación para los gráficos  $u$ , empleando una simbología más general.

#### 3.5.1 Límites de control con un ajuste en la expansión de Cornish-Fisher

Winterbottom (1993) incorporó una modificación simple en los límites de control del gráfico  $c$  mediante la incorporación de un término de ajuste derivado de la expansión cuantílica de Cornish-Fisher. A partir de una variable aleatoria  $Y$  con distribución Poisson de media  $c$ , donde  $\sigma^2 = \mu_3 = c$ , se tiene que el percentil  $\alpha$ -ésimo de la distribución de  $Y$  es;

$$Y_\alpha = c + z_\alpha \sigma^2 + \mu_3 (z_\alpha^2 - 1)/(6\sigma^2) \quad 3.23$$

(Recordar que para la distribución Poisson, el momento central de orden 3,  $\mu_3$ , es igual a la varianza y  $K_3 = \mu_3/\sigma^2$ )

Luego, de 3.25 se tiene que:

$$Y_\alpha = c + z_\alpha \sqrt{c} + (z_\alpha^2 - 1)/6 \quad 3.24$$

Con límites de control  $3\sigma$  para el gráfico  $c$ , dados por:

$$\text{LIC} = c - 3\sqrt{c} + 4/3 \quad 3.25$$

$$\text{LSC} = c + 3\sqrt{c} + 4/3 \quad 3.26$$

Cuando no se conoce  $c$ , se lo puede especificar o remplazar por un valor estimado  $\bar{c}$ .

Para una aplicación a los gráficos  $u$ , Winterbottom consideró una variable aleatoria  $X$  con distribución Poisson con media  $nc$ , donde para una muestra de tamaño  $n$ , la variable aleatoria  $U = X/n$  indica el número de defectos por unidad y tiene media  $c$ , varianza  $\sigma^2 = c/n$ , y tercer momento respecto a  $c$ ,  $\mu_3 = c/n^3$ .

Chen y Cheng encontraron una diferencia de cálculo en  $\mu_3 = c/n^3$  estableciendo que el verdadero valor de  $\mu_3$  era  $\mu_3 = c/n^2$ .

En su trabajo de 1998, revisaron estos cálculos y propusieron un gráfico  $u$  modificado utilizando una nomenclatura más general, basada en los siguientes resultados:

Sea  $U_i = Y_i/n_i$ ,  $\mu_2 = E\{(U_i - \lambda)^2\} = \lambda/n_i$ ,  $\mu_3 = E\{(U_i - \lambda)^3\} = \lambda/n_i^2$ , y sea  $z_\alpha$  el  $\alpha$ -ésimo cuantil de la distribución normal estándar. Luego, el  $\alpha$ -ésimo cuantil de  $U_i$ , denotado por  $u_{i,\alpha}$  es;

$$\begin{aligned} u_{i,\alpha} &= \lambda + z_\alpha \sqrt{\mu_2} + \mu_3(z_\alpha^2 - 1)/(6 \mu_2) \\ &= \lambda + z_\alpha \sqrt{\lambda/n_i} + (z_\alpha^2 - 1)/(6 n_i) \end{aligned} \quad 3.27$$

Luego, para el caso en que  $\lambda$  es conocido y  $z_\alpha = \pm 3$ , los límites de control se obtienen como:

$$\text{LIC} = \lambda - 3\sqrt{\lambda/n_i} + 4/3n_i \quad 3.28$$

$$\text{LSC} = \lambda + 3\sqrt{\lambda/n_i} + 4/3n_i \quad 3.29$$

Para el caso en que  $\lambda$  es desconocido, se define  $\hat{\lambda}_{i-1} = \sum_{j=1}^{i-1} y_j / \sum_{j=1}^{i-1} n_j$ , para  $i = 1, 2, \dots$ . Luego, se grafican los valores  $u_i$  sobre un gráfico con límites de control variables.

### 3.5.2 Límites de control con dos ajustes en la expansión de Cornish-Fisher

Cavalcanti y Cordeiro (2006), generalizaron los resultados obtenidos por Chen y Cheng (1998), incluyendo términos del orden  $n^{-3/2}$  en un gráfico  $u$  mejorado. Ellos también demostraron que con un nuevo término de ajuste a los límites de control se logra una mejora sustancial respecto a la modificación con sólo un término de corrección. Además, esta mejora es equivalente a usar la distribución de Poisson exacta para construir los gráficos.

El gráfico propuesto por estos autores es como sigue:

Sea  $U = Y/n$ , el promedio de no conformidades en una muestra de  $n$  unidades. Los momentos centrales de  $U$  vienen dados por:  $\mu_2 = E\{(U - \lambda)^2\} = \lambda/n$ ,  $\mu_3 = E\{(U - \lambda)^3\} = \lambda/n^2$  y  $\mu_4 = E\{(U - \lambda)^4\} = \lambda(3n\lambda + 1)/n^3$ , y sea  $z_\alpha$  el  $\alpha$ -ésimo cuantil de la distribución normal estándar. Luego, el  $\alpha$ -ésimo cuantil de  $U$ , denotado por  $u_\alpha$  se obtiene de la expansión de Cornish-Fisher con un término adicional a partir del orden  $n^{-1}$ .

$$\frac{u_\alpha - \lambda}{\sqrt{\lambda/n}} = z_\alpha + \frac{1}{6\sqrt{\lambda n}}(z_\alpha^2 - 1) + \frac{1}{24\lambda n}(z_\alpha^3 - 3z_\alpha) - \frac{1}{36\lambda n}(2z_\alpha^3 - 5z_\alpha) \quad 3.30$$

De lo que sigue inmediatamente que;

$$u_\alpha = \lambda + z_\alpha\sqrt{\lambda/n} + (z_\alpha^2 - 1)/(6n) + z_\alpha(1 - z_\alpha^2)/72n\sqrt{\lambda n} \quad 3.31$$

que generaliza los resultados obtenidos por Chen y Cheng (1998).

Finalmente los límites de control variables para graficar  $u_i = y_i/n_i$  donde  $y_i$  indica el número de defectos en la  $i$ -ésima muestra de tamaño  $n_i$ , para el caso en que  $\lambda$  es conocido, están dados por:

$$LIC = \lambda - 3\sqrt{\lambda/n_i} + 4/3n_i - 1/3n_i\sqrt{\lambda n_i} \quad 3.32$$

$$LSC = \lambda + 3\sqrt{\lambda/n_i} + 4/3n_i - 1/3n_i\sqrt{\lambda n_i} \quad 3.33$$

Para el caso en que  $\lambda$  es desconocido, se define  $\hat{\lambda}_{i-1} = \sum_{j=1}^{i-1} y_j / \sum_{j=1}^{i-1} n_j$ , para  $i = 1, 2, \dots$  y se sustituye  $\lambda$  en las Ecuaciones 3.32 y 3.33 por este valor.

Cavalcanti y Cordeiro (2006) mostraron la ventaja de la utilización de estos gráficos modificados mediante la aplicación de un ejemplo tomado de DeVor (1992) para el caso  $\lambda = 1,4$ . Ellos encontraron diferencias con respecto a la propuesta de un sólo término de corrección, especialmente para muestras pequeñas, de tamaño menor a 5. Posteriormente ampliaron sus resultados mediante estudios de simulación.

### 3.6 Comentario final

A partir del estudio realizado considerando la modificación en los límites de control del gráfico  $p$  mediante la incorporación de uno y de dos términos de corrección, se pudo lograr un importante avance en el control de la fracción de no conformes.

Al respecto, y como aporte original de esta investigación, se sugiere una regla a efectos de considerar la elección apropiada de los límites de control para el gráfico  $p$ , a saber:

- ✓ cuando  $np(1 - p) \geq 5$  sin corrección (normal)
- ✓ cuando  $np(1 - p) \geq 0,25$  un término de corrección
- ✓ cuando  $np(1 - p) \geq 0,08$  dos términos de corrección

La ventaja práctica asociada con esta propuesta radica en el hecho de que el usuario puede insertar estos nuevos límites de control en un software estadístico de una manera muy sencilla.

Cabe destacar que la modificación introducida por Cavalcanti y Cordeiro (2006) en los gráficos  $u$  es similar a la efectuada en esta investigación para los gráficos  $p$ .

En ambos casos, se ha demostrado que la incorporación de un nuevo término de corrección en la expansión cuantílica de Cornish-Fisher, ha resultado en grandes beneficios para la utilización de gráficos de control de atributos cuando el número de unidades con defectos o unidades no conformes es muy reducido, situación que se ha tornado habitual actualmente en muchas empresas.

Los resultados obtenidos con esta investigación fueron publicados en el *Journal of Control Engineering Practice*, como se muestra en el Anexo II.

## Capítulo 4

### Gráfico de control para $p$ con muestreo doble

#### 4.1 Introducción

En el Capítulo anterior, se han presentado una serie de alternativas para mejorar la potencia y la sensibilidad de los gráficos  $p$  en proceso de alta calidad. No obstante, aunque los procedimientos propuestos pueden aumentar la exactitud del monitoreo y control de los procesos, todavía carecen de la precisión deseable cuando la proporción de no conformidades  $p$  es muy pequeña y  $n$  no es lo suficientemente grande.

Entre las alternativas propuestas, se presentó un gráfico  $p$  mejorado que, mediante la incorporación de dos términos de ajuste en los límites de control, mostró un beneficio considerable sobre el gráfico  $p$  habitual para atributos. Este nuevo gráfico, basado en la expansión cuantílica de Cornish-Fisher, mostró ser mejor que el tradicional gráfico de control de Shewhart especialmente en el sentido de que permite el monitoreo de valores de la proporción  $p$  más bajos, como es el caso en los procesos de alta calidad.

Sin embargo, el gráfico  $p$  mejorado no es eficiente para detectar pequeños incrementos de los parámetros del proceso. Para estas situaciones, una alternativa a los gráficos de control para atributos con muestreo simple (MS) es la aplicación de gráficos de control con muestreo doble (MD). El muestreo doble es un caso especial de muestreo múltiple, que consiste en la toma de decisiones en dos pasos, en lugar de en un sólo paso, como es lo habitual.

El gráfico de control con muestreo doble (MD) fue propuesto originalmente por Croasdale (1974), para el caso de gráficos de control para variables. En estos primeros gráficos de control con muestreo doble, la información de la primera y segunda muestra se evalúan por separado y la decisión se toma sólo en función de la segunda muestra. Daudin (1992) mejoró el gráfico de control de Croasdale y propuso un gráfico de control con muestreo doble que utiliza la información de ambas muestras en la segunda etapa. El mayor tamaño de la muestra mejora la precisión del gráfico de control dado que utiliza una menor desviación estándar muestral en la estimación de los límites de control. El gráfico de control de Daudin optimiza el tamaño de muestra esperado. Irianto y Shinozaki (1998) maximizaron la potencia del gráfico de control para determinar los límites de control del gráfico. He *et al.* (2002) y Costa y Claro (2007) han hecho otros desarrollos en los gráficos de control con muestreo doble para el caso de gráficos de control para variables.

Recientemente se han realizado otras investigaciones tendientes a mejorar la efectividad de los gráficos de control de atributos, especialmente para los gráficos  $np$ . Wu *et al.* (2006) desarrollaron un algoritmo para la optimización del diseño del gráfico de control  $np$  considerando una inspección del 100%. Wu y Wang (2007) propusieron un gráfico de control  $np$  con doble inspección. La primera inspección decide el estado del proceso (bajo control o fuera de control) según el número de unidades no conformes encontradas en la muestra y, en la segunda inspección, el gráfico propuesto controla la ubicación de una unidad no conforme particular en la muestra.

Rodrigues *et al.* (2011), propusieron un muestreo en dos etapas o muestreo doble (MD) empleando los gráficos de control de Shewhart en la primera etapa. Este es otro método utilizado para mejorar el desempeño del gráfico  $p$  tradicional de Shewhart sin incrementar el número promedio (bajo control) de unidades inspeccionadas. Durante la primera etapa, son inspeccionadas una o más unidades de la muestra y, dependiendo del resultado, el muestreo se interrumpe o pasa a la segunda etapa, donde las unidades restantes de la muestra son inspeccionadas.

El propósito del muestreo doble radica en reducir el número de unidades a ser inspeccionadas. La eficacia de este plan consiste en detectar cambios pequeños o moderados en la proporción de unidades no conformes en el proceso. Generalmente en situaciones como ésta es necesario tomar muestras grandes, lo que pueden ser indeseable o inviable en algunas situaciones prácticas.

Un plan de muestreo doble tiene dos posibles ventajas sobre un plan de muestreo simple. En primer lugar, puede reducir la cantidad total de inspección. La primera muestra tiene un tamaño menor respecto de lo que se pide bajo un plan de muestreo simple y por consiguiente, en todos los casos en que un lote sea aceptado o rechazado en la primera muestra, puede haber un ahorro considerable en la inspección total. También es posible rechazar un lote sin inspeccionar completamente la segunda muestra. En segundo lugar, un plan de muestreo doble tiene la ventaja de dar al lote una segunda oportunidad. A algunas personas, especialmente a quienes controlan el proceso, puede parecerles injusto rechazar un lote sobre la base de una sola muestra. Es decir, el muestreo doble le da a un proceso una nueva oportunidad antes de tomar una decisión.

En la práctica, especialmente en la industria, el caso de emplear muestras pequeñas para monitorear los procesos mediante un gráfico de control es más común que el caso de emplear muestras grandes y por lo tanto, es necesario desarrollar metodologías de control de procesos que tengan en cuenta esta situación.

El procedimiento en dos etapas propuesto por Rodrigues *et al.* (2011) para monitorear la proporción de no conformes en procesos de alta calidad, está basado en la primera etapa en la utilización del gráfico  $p$  de Shewhart que es válido sólo para muestras grandes.

En este Capítulo, se extienden los resultados obtenidos por Rodrigues *et al.* (2011), al caso de procesos de alta calidad. Esta nueva versión del MD está basada en la utilización del gráfico  $p$  mejorado en lugar del gráfico tradicional de Shewhart. De esta manera, mediante la corrección cuantílica (límite superior de control) de Cornish-Fisher la aproximación binomial-normal es válida para muestras chicas lo que permite que el procedimiento de MD pueda ser empleado en este tipo de situaciones.

## 4.2 Procedimiento del gráfico de control con muestreo doble.

El gráfico de control con muestreo doble (MD) tiene por objeto mejorar la capacidad de detectar cualquier pequeño cambio de condición de fuera de control mediante la observación de una segunda muestra sin interrumpir el muestreo. Además permite preservar las ventajas del gráfico de control para  $p$  e incrementar la capacidad de detectar desplazamientos de la fracción de no conformes del proceso, reduciendo al mismo tiempo el tamaño de la muestra. El procedimiento de control para el MD es similar al del MS por lo que el plan de muestreo simple presentado en el Capítulo 2, puede ser extendido al esquema de muestreo doble.

El procedimiento del gráfico es el siguiente:

Sea  $p$  la proporción de unidades no conformes y  $p_0$  la proporción de unidades no conformes cuando el proceso está bajo control, es decir, cuando  $p = p_0$ . Si se presentan causas asignables en el proceso, habrá un incremento en el valor de  $p$  (es decir,  $p = p_1$ , con  $p_1 > p_0$ ). Sólo este caso será considerado, por lo que el gráfico se define sin límite de control inferior debido a que en procesos con baja proporción de no conformes el límite inferior es nulo.

El gráfico  $np$  con MD utiliza cinco parámetros: el tamaño de la primera muestra ( $n_1$ ), el número de aceptación para la primera muestra (límite de advertencia, LA), el número de rechazo para la primera muestra  $LSC_1$ , el tamaño de la segunda muestra ( $n_2$ ) y el número de aceptación (no rechazo) para la segunda etapa de muestreo ( $LSC_2$ ).

Luego, dados  $n_1$ ,  $n_2$ , LA,  $LSC_1$  y  $LSC_2$ , es posible definir un plan de control como;

- 1.- Tomar una muestra de  $n_1$  unidades del proceso
- 2.- Sea  $d_1$  el número de unidades no conformes encontrados en esta muestra (de tamaño  $n_1$ )  
En este caso, la decisión depende de  $d_1$ ;
- 2.1.- Si  $d_1 < LA$ , el proceso es considerado bajo control y se continúa operando con el esquema de control previsto.
- 2.2.- Si  $d_1 > LSC_1$ , el proceso se considera fuera de control debiendo tomar acciones correctivas.

2.3.- Si  $LA < d_1 \leq LSC_1$ , una muestra adicional de  $n_2$  unidades debe ser tomada inmediatamente.

Cuando se da la situación del punto 2.3,

- 3.- Tomar una segunda muestra de  $n_2$  unidades del proceso.
- 4.- Sea  $d_2$  el número de unidades no conformes encontrados en esta segunda muestra (de tamaño  $n_2$ )
- 5.- Sea  $(d_1 + d_2)$ , la suma de unidades no conformes encontrados en las dos muestras (con tamaños  $n_1$  y  $n_2$ ).

En este caso, la decisión depende de  $(d_1 + d_2)$

- 5.1.- Si  $(d_1 + d_2) < LSC_2$ , el proceso es considerado bajo control y se continúa operando con el esquema de control previsto.
- 5.2.- Si  $(d_1 + d_2) > LSC_2$ , el proceso se considera fuera de control debiendo tomar acciones correctivas.
- 6.- En intervalos de muestreo prefijados (cada hora, por ejemplo) retornar a la primera fase del esquema de MD y tomar una nueva muestra de tamaño  $n_1$ .

Estas acciones se muestran en la Figura 4.1

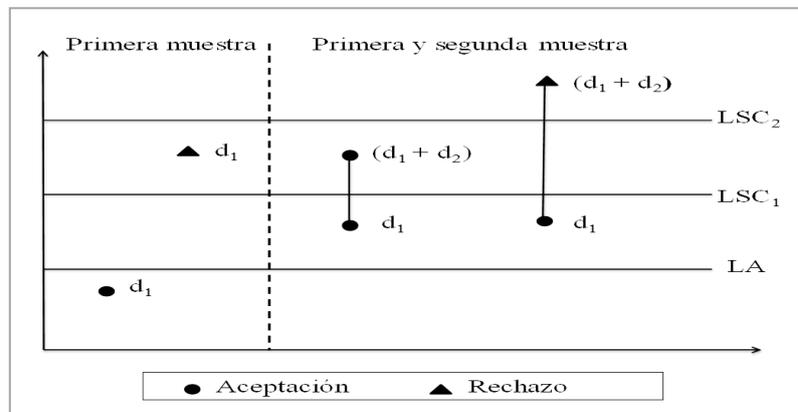


Figura 4.1: Procedimiento del gráfico de control con MD

Debe tenerse en cuenta que el número de unidades no conformes en ambas muestras,  $d_1$  y  $d_2$ , son siempre valores enteros, por lo que resulta apropiado que los límites de advertencia y de control sean expresados mediante valores fraccionarios. Cuando se emplean límites fraccionarios, el usuario no dudará en determinar si un número de unidades no conformes se encuentra dentro o fuera de los límites del control. De esta manera, los límites de advertencia y de control fraccionarios se pueden definir de la siguiente manera:

$$LA = [LA] + 0.5 ; \quad LSC_1 = [LSC_1] - 0.5 ; \quad LSC_2 = [LSC_2] + 0.5$$

Donde  $[LA]$  representa el “piso” de su argumento (es decir, el mayor entero menor o igual al argumento) del límite de advertencia,  $[LSC_1]$  representa el "techo" de su argumento (es decir, el menor número entero mayor o igual que el argumento) del límite de control de la primera muestra y  $[LSC_2]$  representa el mayor entero menor o igual al argumento para el límite de control con la segunda muestra.

Siguiendo el mismo criterio, para el gráfico con muestreo simple,  $LSC_0 = [LSC] + 0.5$ , donde  $[LSC]$  representa el mayor entero menor o igual al argumento del límite superior de control.

### 4.3 Medida de desempeño del gráfico

En el área de calidad y principalmente en la industria de la producción, una forma de evaluar las decisiones con respecto a la eficacia de un gráfico control es a través de la longitud media de corrida (ARL). Como ya fue definido, el ARL es el número promedio de puntos que deben ser graficados antes de que un punto indique una condición de fuera de control. En particular,  $ARL_0$  es la longitud promedio de corrida para un proceso en estado de control y  $ARL_1$  es la longitud promedio de corrida para un proceso fuera de control.

Se conoce que incluso cuando un proceso está bajo control (es decir, cuando no hay causas especiales presentes en el sistema,  $p = p_0$ ), hay aproximadamente una probabilidad de 0,0027 de que un punto exceda el límite de control superior  $3\sigma$ . Incluso en un gráfico de control apropiadamente construido, puede ser posible que se observe la presencia de una causa especial sin que necesariamente ella haya ocurrido.

Por otro lado, si se produce una causa especial, ésta puede no ser de magnitud suficiente como para que el gráfico determine una condición inmediata de alarma. Si se produce una causa especial, uno puede describir esa causa midiendo el cambio en la fracción de unidades no conformes en el proceso en cuestión ( $p_1 > p_0$ ). Cuando se cuantifican los cambios, es posible determinar el ARL fuera de control para el gráfico ( $ARL_1$ ).

Sean,  $ARL_0$  y  $ARL_1$  las medidas de rendimiento de los gráficos de control, con  $ARL_0 = 1/\alpha$ , donde  $\alpha$  es la probabilidad de falsa alarma y  $ARL_1 = 1/1 - \beta$ , donde  $\beta$  es la probabilidad de no detectar una causa especial en la muestra. En los gráficos de control con muestreo simple, estas cantidades se obtienen directamente utilizando las ecuaciones 2.10 y 2.11 del Capítulo 2.

En MD, estas medidas de desempeño se obtienen de manera diferente. Para el gráfico  $np$  con MD, después de cada muestreo, la probabilidad de que el gráfico indique que el proceso está bajo control, está dada por:

$$P_T = P_a^I + P_a^{II} \quad 4.1.$$

donde  $P_a^I$  indica la probabilidad de que el valor del estadístico en la muestra esté por debajo del límite de advertencia (LA) en la primera etapa del esquema de MD y  $P_a^{II}$  indica la probabilidad de que se necesite tomar una segunda muestra y que el número total de unidades no conformes en las dos muestras sea menor que el límite de control de la segunda etapa (LSC<sub>2</sub>).

$P_a^I$  y  $P_a^{II}$  están dados por:

$$P_a^I = Pr(d_1 \leq [LA], n_1, p_0) = \sum_{d_1=0}^{[LA]} \frac{n_1!}{d_1! (n_1 - d_1)!} p_0^{d_1} (1 - p_0)^{n_1 - d_1} \quad 4.2.$$

Donde  $P_a^I$  es la probabilidad de aceptar el lote con la primera muestra.

$$P_a^{II} = Pr([LA] < d_1 < [LSC_1]) \cap (d_1 + d_2 \leq [LSC_2])$$

$$P_a^{II} = \sum_{d_1=[LA]+1}^{[LSC_1]-1} \left[ \frac{n_1!}{d_1! (n_1 - d_1)!} p_0^{d_1} (1 - p_0)^{n_1 - d_1} \right] \times \left( \sum_{d_2=0}^{[LSC_2]-d_1} \frac{n_2!}{d_2! (n_2 - d_2)!} p_0^{d_2} (1 - p_0)^{n_2 - d_2} \right) \quad 4.3.$$

y  $P_a^{II}$  es la probabilidad de aceptar (no rechazar) el lote cuando una segunda muestra debe ser tomada

De esta manera, es posible calcular  $\alpha$ ,  $\beta$ ,  $ARL_0$  y  $ARL_1$  como:

$$\text{Cuando } p = p_0, \quad \alpha = 1 - P_T, \quad ARL_0 = 1 / (1 - P_T), \quad y \quad 4.4.$$

$$\text{Cuando } p = p_1, \quad \beta = P_T, \quad ARL_1 = 1 / 1 - P_T$$

No menos importante es determinar el tamaño promedio de muestra que termina siendo considerado. En el muestreo simple, el tamaño de la muestra total inspeccionada es siempre constante. En el muestreo doble este tamaño variará dependiendo de si una segunda muestra es necesaria o no. Con una inspección completa de la segunda muestra, el tamaño promedio de la muestra es igual al tamaño de la primera muestra por la probabilidad de que sólo se necesite una muestra, más el tamaño de la muestra combinada por la probabilidad de que la segunda muestra sea necesaria.

Cuando  $p = p_0$ , la probabilidad de tomar una decisión en la primera muestra es  $P_1$ , esto es;

$P_1 = Pr$  (el lote sea aceptado en la primera muestra) +  $Pr$  (el lote sea rechazado en la primera muestra)

$$P_1 = Pr(d_1 \leq [LA], n_1, p_0) + Pr(d_1 > [LSC_1], n_1, p_0) \quad 4.5.$$

Luego, la probabilidad de tener que tomar una segunda muestra es  $1 - P_1$ , donde;

$$1 - P_1 = Pr([LA] < d_1 < [LSC_1]) \quad 4.6.$$

Por lo tanto, una formula general para el tamaño promedio de la muestra (TPM) en el muestreo doble, asumiendo inspección completa de la segunda muestra es;

$$TPM = n_1 P_1 + (n_1 + n_2)(1 - P_1) = n_1 + n_2(1 - P_1) \quad 4.7.$$

Siendo  $P_1$  la probabilidad de tomar una decisión en la primera muestra.

En la práctica, la inspección usualmente se realiza empleando ambas muestras y el lote es rechazado tan pronto como el número de unidades no conformes en la muestra combinada, supera el segundo número de aceptación  $LSC_2$ . De esta manera, es posible rechazar un lote sin una inspección completa de la segunda muestra (esto se conoce como cercenar o restringir la segunda muestra). Por lo tanto, el muestreo doble frecuentemente puede llevar a reducir los costos generales de inspección.

#### 4.4 Ventaja del gráfico $np$ mejorado

En esta sección se muestra la ventaja de emplear el gráfico  $np$  mejorado respecto al gráfico  $np$  basado en la distribución normal, para el caso  $n$  y  $p$  chicos. La comparación entre los dos procedimientos se realiza mediante la evaluación del  $ARL_0$  que indica la cantidad de puntos a graficar, antes de dar una señal de alarma, cuando el proceso está bajo control.

Como un ejemplo, supóngase que se toma  $p_0 = 0,005, 0,01$  y  $0,02$  con  $np_0 = 0,2, 0,3, 0,4$  y  $0,5$  (valores que más tarde se utilizan en las Tablas 4.2 a 4.5) y  $ARL_{0min} = 370,4$ . Recordar que en el MS el problema de optimización del gráfico se reduce a encontrar el menor valor del LSC para el cual  $ARL_0 \geq ARL_{0min}$ .

En la Tabla 4.1 se observa que los valores de  $ARL_0$  para el gráfico mejorado son mayores que los mismos valores para el gráfico de Shewhart. En esta situación el gráfico basado en la distribución normal muestra señales de falsa alarma en periodos más cortos que el prefijado lo que se conoce como “excesivas señales de alarma”. Nótese también que el límite superior de control del gráfico de Shewhart para  $np_0 = 0,2$  y  $0,3$ , es igual a 1 (por ser el mayor valor entero del LSC). Esto

indica que el proceso será aceptado (no rechazado) sólo en caso en que en la muestra se observe una o ninguna unidad no conforme. Esta limitación condiciona además el procedimiento de MD que deja de ser una opción posible.

Tabla 4.1: Diseño comparativo del gráfico  $np$  con MS, con  $ARL_{0min} = 370,4$

$p_0$	$n$	$np_0$	Gráfico $np$ mejorado con MS		Gráfico $np$ de Shewhart con MS		% ganancia
			LSC	$ARL_0$	LSC	$ARL_0$	
0.005	40	0,2	2,109	929,88	1,538	58,16	93,75
	60	0,3	2,647	289,17	1,939	27,36	90,54
	80	0,4	3,083	1369,21	2,293	129,67	90,53
	100	0,5	3,462	597,63	2,616	70,91	88,13
0.01	20	0,2	2,088	996,44	1,535	59,31	94,05
	30	0,3	2,627	301,41	1,935	27,66	90,82
	40	0,4	3,062	1458,56	2,288	133,38	90,86
	50	0,5	3,441	626,50	2,611	72,37	88,45
0.02	10	0,2	2,048	1157,53	1,528	61,81	94,66
	15	0,3	2,586	329,02	1,927	28,30	91,34
	20	0,4	3,021	1667,56	2,278	141,47	91,52
	25	0,5	3,399	691,62	2,600	75,51	89,08

En la última columna de la tabla se muestra el porcentaje de ganancia del  $ARL_0$  cuando se utiliza el gráfico  $np$  mejorado respecto al gráfico  $np$  de Shewhart para el caso del MS. Este porcentaje se obtuvo empleando la siguiente relación:

$$\% \text{ ganancia} = \frac{ARL_{0 \text{ mejorado}} - ARL_{0 \text{ Shewhart}}}{ARL_{0 \text{ mejorado}}} \times 100 \quad 4.8.$$

A modo de ejemplo, cuando  $np_0 = 0,5$ ,  $p_0 = 0,005$  y  $n = 100$ , con  $ARL_{0min} = 370,4$ , el porcentaje de ganancia =  $((597,63 - 70,91) / 597,63) \times 100\% = 88,13\%$ . Este resultado indica que cuando se utiliza el gráfico  $np$  mejorado en lugar del gráfico  $np$  basado en la distribución normal, para este ejemplo en particular, se obtiene una ganancia de aproximadamente un 88% en términos de la cantidad promedio de puntos a graficar antes de que detecte una señal de falsa alarma.

#### 4.5 Diseño y aplicación del muestreo doble al gráfico $np$ mejorado

Ahora es posible definir el diseño óptimo del gráfico  $np$  con MD y el desempeño comparativo de este gráfico con el gráfico  $np$  con MS. El diseño del gráfico  $np$  con MD se basa en ciertas condiciones de optimización como fueron establecidas por Rodrigues *et al.* (2011).

Sea  $TPM_1$  el tamaño promedio de muestra para un proceso fuera de control, es decir cuando  $p = p_1$ , y  $TPM_0$  el tamaño promedio de muestra para un proceso bajo control con  $p = p_0$ . Para un

desplazamiento determinado de  $p$ , la función objetivo consiste en minimizar  $\beta$  o equivalentemente minimizar  $ARL_1$  sujeto a ciertas restricciones sobre  $TPM_0$  y  $ARL_0$ .

Formalmente, el problema de optimización del gráfico  $np$  con MD es el siguiente:

- a. Minimizar  $ARL_1$  (o minimizar el número de muestras hasta la detección de una unidad fuera de control en el proceso), sujeto a:
- b.  $TPM_0 \leq n$  (el tamaño promedio de muestra en MD con  $p = p_0$ , debe ser menor o igual al tamaño de la muestra en el muestreo simple, MS) y
- c.  $ARL_0 \geq ARL_{0min}$

Donde  $n$  y  $ARL_{0min}$  son valores especificados.

Las expresiones para calcular  $ARL_0$ ,  $ARL_1$  y  $TPM_0$  están dadas en las Ecuaciones 4.4 y 4.7.

Con el fin de permitir la comparación entre los gráficos  $np$  con MD y MS bajo las mismas condiciones, se imponen ciertas restricciones sobre  $TPM_0$  y  $ARL_0$  bajo control. Para un número de valores especificados de estas restricciones, es posible comparar los valores ARL de ambos gráficos para un rango de desplazamientos ascendentes en los valores de  $p$ .

Los parámetros asociados al gráfico con MS son el tamaño especificado de la muestra  $n$  y el límite superior de control LSC, dado que el gráfico se define sin límite inferior de control. En el gráfico con MD, los valores de  $p_0$ ,  $p_1$ ,  $n$  y  $ARL_{0min}$ , son valores dados. Además, la proporción de no conformidades  $p_1$ , se obtiene de la relación  $\gamma = p_1/p_0$ .

Para resolver el problema de optimización, se confeccionó un programa bajo plataforma de R, basado en el programa original en Matlab<sup>TM</sup> que se encuentra en la tesis doctoral de Rodrigues (2005).

Para comprender el procedimiento del gráfico, supóngase un proyecto para procesos de alta calidad con una proporción de no conformidades baja y muestras chicas. Los parámetros de entrada considerados son:  $p_0 = 0,005, 0,01, 0,02$ ;  $\gamma = 1,5, 2,0$ ;  $n = 10, 15, 20, 25, 30, 40, 50, 60, 80, 100$ , y  $ARL_{0min} = 200$  y  $370,4$ . Un  $ARL_{0min} = 200$  corresponde a un riesgo de falsa alarma de  $0,005$  ( $z_\alpha = 2.807$ ) y un  $ARL_{0min} = 370,4$  corresponde a un riesgo de falsa alarma de  $0,0027$  ( $z_\alpha = 3$ ).

Un  $ARL_{0min} = 370,4$  es considerado un valor estándar para  $ARL_0$  en los gráficos de control, pero este valor puede conducir a tamaños de muestra muy grandes si se requiere que el gráfico tenga un poder de detección razonable para una situación de fuera de control.

En las Tablas 4.2 a 4.5, es posible ver los proyectos óptimos para los gráficos  $np$  con MS y MD y las medidas de rendimiento para todas las combinaciones de  $\gamma$ ,  $ARL_{0min}$ ,  $p_0$  y  $TPM$ , previamente

establecidas. Las Tablas 4.2 y 4.3 muestran los resultados para  $ARL_{0min} = 370,4$ , con  $\gamma = 1,5$ , y  $2,0$ , mientras que las Tablas 4.4 y 4.5 muestran los resultados para  $ARL_{0min} = 200$  con  $\gamma = 1,5$  y  $2,0$ . También es posible observar los valores de  $ARL_0$  y  $ARL_1$  así como los parámetros de diseño óptimo para cada uno de los gráficos considerados.

Tabla 4.2: Diseño óptimo para gráficos  $np$  con MS y MD con  $ARL_{0min} = 370,4$  y  $\gamma = 1,5$

		Gráfico de control $np$ con muestreo simple						Gráfico de control $np$ con muestreo doble						
$p_0$	$p_1$	$n$	LSC	LSC <sub>0</sub>	ARL <sub>0</sub>	ARL <sub>1</sub>	LA	LSC <sub>1</sub>	LSC <sub>2</sub>	$n_1$	$n_2$	TPM	ARL <sub>0</sub>	ARL <sub>1</sub>
0.5%	0.75%	40	2.109	2.5	929.88	295.17	1.5	2.5	4.5	34	162	35.94	803.41	193.22
		60	2.647	2.5	<b>289.17</b>	95.21	1.5	2.5	4.5	42	203	45.57	370.98	91.48
				3.5	4102.47	905.70	1.5	2.5	4.5	42	203	45.57	370.98	91.48
		80	3.083	3.5	1369.21	314.31	1.5	3.5	5.5	64	271	75.05	371.19	64.00
		100	3.462	3.5	597.63	142.60	1.5	3.5	6.5	74	352	92.57	372.43	55.45
1%	1.5%	20	2.089	2.5	996.44	314.66	1.5	2.5	4.5	18	81	19.06	738.32	177.41
		30	2.627	2.5	<b>301.41</b>	98.75	1.5	2.5	4.5	21	104	22.80	374.61	91.30
				3.5	4492.41	984.55	1.5	2.5	4.5	22	96	23.81	370.48	91.50
		40	3.062	3.5	1458.56	332.47	1.5	3.5	5.5	32	137	37.53	371.55	63.67
		50	3.441	3.5	626.50	148.47	1.5	3.5	6.5	37	178	46.32	372.23	54.92
2%	3%	10	2.048	2.5	1157.53	361.67	1.5	2.5	4.5	9	37	9.46	923.44	218.99
		15	2.586	2.5	<b>329.02</b>	106.71	1.5	2.5	4.5	11	51	11.93	373.56	89.88
				3.5	5464.31	1179.8	1.5	2.5	4.5	11	51	11.93	373.56	89.88
		20	3.021	3.5	1667.56	374.70	1.5	3.5	5.5	16	70	18.77	372.28	63.03
		25	3.399	3.5	691.62	161.66	1.5	3.5	6.5	19	89	23.82	371.88	53.92

Tabla 4.3: Diseño óptimo para gráficos  $np$  con MS y MD con  $ARL_{0min} = 370,4$  y  $\gamma = 2,0$

		Gráfico de control $np$ con muestreo simple						Gráfico de control $np$ con muestreo doble						
$p_0$	$p_1$	$n$	LSC	LSC <sub>0</sub>	ARL <sub>0</sub>	ARL <sub>1</sub>	LA	LSC <sub>1</sub>	LSC <sub>2</sub>	$n_1$	$n_2$	TPM	ARL <sub>0</sub>	ARL <sub>1</sub>
0.5%	1%	40	2.109	2.5	929.88	133.38	1.5	2.5	4.5	34	162	35.94	803.41	72.74
		60	2.647	2.5	<b>289.17</b>	44.60	1.5	2.5	4.5	42	203	45.57	370.98	36.13
				3.5	4102.47	320.17	1.5	2.5	4.5	43	194	46.56	370.91	36.33
		80	3.083	3.5	1369.21	115.48	1.5	3.5	5.5	65	267	76.20	372.08	21.42
		100	3.462	3.5	597.63	54.42	1.5	3.5	6.5	78	336	97.43	372.29	17.21
1%	2%	20	2.089	2.5	996.44	141.47	1.5	2.5	4.5	18	81	19.06	738.32	66.70
		30	2.627	2.5	<b>301.41</b>	46.05	1.5	2.5	4.5	21	104	22.80	374.61	35.87
				3.5	4492.41	345.60	1.5	2.5	4.5	22	96	23.81	370.50	35.88
		40	3.062	3.5	1458.56	121.35	1.5	3.5	5.5	33	133	38.68	373.58	21.19
		50	3.441	3.5	626.50	56.31	1.5	3.5	6.5	39	170	48.77	372.53	16.97
2%	4%	10	2.048	2.5	1157.53	160.93	1.5	2.5	4.5	9	37	9.46	923.44	80.50
		15	2.586	2.5	<b>329.02</b>	49.28	1.5	2.5	4.5	11	51	11.93	373.56	34.88
				3.5	5464.31	408.21	1.5	2.5	4.5	11	51	11.93	373.56	34.88
		20	3.021	3.5	1667.56	134.90	1.5	3.5	5.5	17	66	19.92	377.35	20.77
		25	3.399	3.5	691.62	60.53	1.5	3.5	6.5	19	89	23.82	371.88	16.55

Tabla 4.4: Diseño óptimo para gráficos  $np$  con MS y MD con  $ARL_{0min} = 200$  y  $\gamma = 1,5$

		Gráfico de control $np$ con muestreo simple						Gráfico de control $np$ con muestreo doble						
$p_0$	$p_1$	$n$	LSC	LSC <sub>0</sub>	ARL <sub>0</sub>	ARL <sub>1</sub>	LA	LSC <sub>1</sub>	LSC <sub>2</sub>	$n_1$	$n_2$	TPM	ARL <sub>0</sub>	ARL <sub>1</sub>
0.5%	0.75%	40	2.109	2.5	929.88	295.17	1.5	2.5	4.5	34	162	35.94	803.41	193.22
		60	2.477	2.5	289.17	95.21	1.5	2.5	4.5	50	242	55.83	200.04	51.35
		80	2.880	2.5	<b>129.67</b>	44.26	1.5	2.5	4.5	56	192	61.63	200.95	53.09
				3.5	1369.21	314.31	1.5	3.5	5.5	66	324	79.96	225.64	42.07
100	3.234	3.5	597.63	142.60	1.5	3.5	5.5	79	290	96.14	200.25	37.01		
1%	1.5%	20	1.968	2.5	996.44	314.66	1.5	2.5	4.5	18	81	19.06	738.32	177.41
		30	2.460	2.5	301.41	98.75	1.5	2.5	4.5	25	118	27.81	212.94	54.04
		40	2.863	2.5	<b>133.38</b>	45.31	1.5	2.5	4.5	28	99	30.88	201.04	52.47
				3.5	1458.56	332.47	1.5	3.5	5.5	33	162	39.91	230.68	42.63
50	3.216	3.5	626.50	148.47	1.5	3.5	5.5	39	148	47.50	200.76	36.93		
2%	3%	10	1.935	2.5	1157.53	361.67	1.5	2.5	4.5	9	37	9.46	923.44	218.99
		15	2.426	2.5	329.02	106.71	1.5	2.5	4.5	13	56	14.40	221.64	55.49
		20	2.828	2.5	<b>141.47</b>	47.60	1.5	2.5	4.5	14	52	15.48	203.14	51.83
				3.5	1657.56	374.70	1.5	3.5	5.5	16	81	19.21	257.64	46.53
25	3.181	3.5	691.62	161.66	1.5	3.5	5.5	19	77	23.17	202.04	36.86		

Tabla 4.5: Diseño óptimo para gráficos  $np$  con MS y MD con  $ARL_{0min} = 200$  y  $\gamma = 2,0$

		Gráfico de control $np$ con muestreo simple						Gráfico de control $np$ con muestreo doble						
$p_0$	$p_1$	$n$	LSC	LSC <sub>0</sub>	ARL <sub>0</sub>	ARL <sub>1</sub>	LA	LSC <sub>1</sub>	LSC <sub>2</sub>	$n_1$	$n_2$	TPM	ARL <sub>0</sub>	ARL <sub>1</sub>
0.5%	1%	40	2.109	2.5	929.88	133.38	1.5	2.5	4.5	34	162	35.94	803.41	72.74
		60	2.477	2.5	289.17	44.60	1.5	2.5	4.5	50	242	55.83	200.04	21.37
		80	2.880	2.5	<b>129.67</b>	21.48	1.5	2.5	4.5	54	209	59.76	200.11	21.51
				3.5	1369.21	115.48	1.5	3.5	5.5	66	324	79.96	225.64	15.26
100	3.234	3.5	597.63	54.42	1.5	3.5	5.5	81	283	98.47	200.52	13.14		
1%	2%	20	1.968	2.5	996.44	141.47	1.5	2.5	4.5	18	81	19.06	738.32	66.70
		30	2.460	2.5	301.41	46.05	1.5	2.5	4.5	25	118	27.81	212.94	22.24
		40	2.863	2.5	<b>133.38</b>	21.90	1.5	2.5	4.5	28	99	30.88	201.04	21.52
				3.5	1458.56	121.35	1.5	3.5	5.5	33	162	39.91	230.68	15.36
50	3.216	3.5	626.50	56.31	1.5	3.5	5.5	41	141	49.84	201.34	13.02		
2%	4%	10	1.935	2.5	1157.53	160.93	1.5	2.5	4.5	9	37	9.46	923.44	80.50
		15	2.426	2.5	329.02	49.28	1.5	2.5	4.5	13	56	14.40	221.64	22.45
		20	2.828	2.5	<b>141.47</b>	22.80	1.5	2.5	4.5	14	52	15.48	203.14	21.04
				3.5	1657.56	134.90	1.5	3.5	5.5	16	81	19.21	257.64	16.48
25	3.181	3.5	691.62	60.53	1.5	3.5	5.5	20	73	24.33	205.27	13.04		

El análisis muestra que algunas combinaciones de parámetros considerados en los proyectos óptimos producen grandes valores  $ARL_1$ , valores que son inaceptables en la práctica. Sin embargo, estos proyectos son de interés para comparar el desempeño de los gráficos  $np$  con MS y MD.

Debe notarse que para algunos valores del límite superior de control en el gráfico con MS, la acción produce dos posibilidades para el análisis (coloreados en gris). Esto se debe a que el menor valor entero del límite de control no cumple con una de las restricciones de optimización establecidas, dado que el  $ARL_{0min}$  es inferior al valor prefijado.

En los gráficos de control de atributos, los límites de control son valores fraccionarios, pero la estadística graficada es un valor entero. Por ejemplo, si el  $LSC = 2,880$ , esto significa que si en la muestra se encuentran más de dos artículos defectuosos, el proceso es considerado fuera de control. Lo mismo pasaría si  $LSC = 2,043$  puesto que la variable binomial considera solamente el valor entero del número. Obviamente los límites de control pueden ser "redondeados" a valores enteros (como se hace aquí), pero cuanto mayor es el valor del límite superior de control (para un mismo tamaño de muestra y valor de  $p$ ), mayor será el  $ARL_1$  o número de muestras hasta la detección de una unidad fuera de control en el proceso. En las tablas anteriores se colocan las dos posibilidades, la primera que es la que corresponde pero que no cumple con las restricciones, y la segunda que es la que satisface las restricciones. Sin embargo, aunque la primera no cumple con las restricciones en el muestreo simple, es la que tiene el mejor rendimiento en el muestreo doble.

Para ilustrar el proceso de optimización, se consideran los datos de la Tabla 4.5, con:

$$\gamma = 2,0 ; \alpha_{\max} = 0,005 \quad \text{o} \quad ARL_0 = 200; \quad p_0 = 0,5\% , \quad [LSC] = 2 \quad \text{y} \quad TPM_{\max} = 60$$

(Segunda fila de la tabla)

El diseño óptimo del gráfico  $np$  con MD es (50, 242, 1,5, 2,5, 4,5) dado que  $n_1 = 50$ ,  $n_2 = 242$ ,  $[LA] = 1$ ,  $[LSC_1] = 3$  y  $[LSC_2] = 4$ . Con estos parámetros,  $TPM = 55.83$  y las medidas de desempeño son  $ARL_0 = 200,04$  y  $ARL_1 = 21,37$ .

El proyecto para el gráfico de control  $np$  mejorado con MS que logra las mismas condiciones con;  $n \leq 60$  y  $\alpha = 0,005$  es el gráfico de control con MS (60; 2,5) donde  $n = 60$ ,  $[LSC] = 2$ ,  $ARL_0 = 289,17$  y  $ARL_1 = 44,60$ .

Luego, el gráfico  $np$  con muestreo doble y el gráfico  $np$  con muestreo simple son comparables en el sentido de que cumplen las mismas restricciones respecto a TPM y  $ARL_0$  pero el gráfico  $np$  con MD tiene menor  $ARL_1$ .

En la siguiente sección se muestra un ejemplo empleando este plan de muestreo doble.

#### 4.6 Aplicación del MD a un proceso de pintura de autopartes.

El problema a analizar está referido a la adhesión de un sellador de pintura utilizado en piezas de automóvil. El proceso consiste en pintar las piezas con este sellador y luego aplicar la pintura final. Una pieza se considera no conforme cuando no consigue una buena adherencia de la pintura final. Para el control, se toma una muestra de 60 piezas/hora de producción y se considera el número de piezas con mala adherencia de la pintura final. A causa de que el muestreo es costoso, el Departamento de Calidad decide implementar el muestreo doble con el fin de reducir el número de muestras a ser inspeccionadas.

En la Tabla 4.6 se muestra el resultado de la inspección inicial del proceso o Fase I (sobre Fase I se puede ampliar en Chakraborti, *et al.*, 2008), realizado sobre 12 muestras de tamaño 60. La proporción de no conformidades estimada fue de 0,005 Posteriormente se agregaron 12 muestras en las que la proporción de no conformidades del proceso se incrementó a 0,01 (claramente esto es sólo una parte del control, dado que el número de muestras en Fase I es siempre mucho mayor).

Considérese nuevamente el proyecto óptimo para los datos de la Tabla 4.5, para el gráfico con MS (60; 2,5) y con diseño óptimo del gráfico  $np$  con MD (50, 242, 1,5, 2,5 y 4,5).

Para llevar a cabo el plan de muestreo doble, se procede de la siguiente manera; se toma una primera muestra de tamaño  $n_1 = 50$  piezas. Si el número de piezas no conformes en la muestra ( $d_1$ ) es menor que  $LA = 1,5$  ( $\hat{o} \leq 1$ ), el proceso es considerado bajo control y se continua con el esquema de control como venía operando. Si  $d_1 > 2,5$  ( $\hat{o} \geq 3$ ) se considera que el proceso está fuera de control. Si  $1,5 < d_1 < 2,5$  ó  $d_1 = 2$ , debe tomarse inmediatamente una muestra adicional de  $n_2 = 242$  piezas. En esta muestra, se cuenta el número de piezas no conformes  $d_2$ , y si  $(d_1 + d_2)$  es menor que  $LSC_2 = 4,5$  ( $\hat{o} \leq 4$ ) el proceso se considera bajo control y se continua con el esquema de control como venía operando. Pero, si  $(d_1 + d_2)$  es mayor que  $LSC_2 = 4,5$  ( $\hat{o} \geq 5$ ) el proceso se considera fuera de control por lo que debe tomarse alguna acción correctiva. Este procedimiento se encuentra en la Tabla 4.6 y graficado en la Figura 4.2.

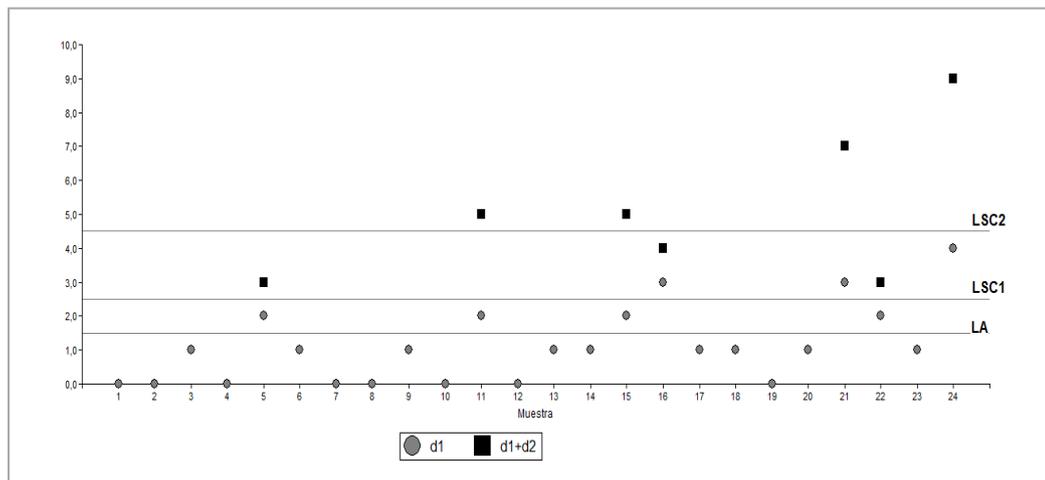


Figura 4.2: Gráfico de control  $np$  para los valores de la Tabla 4.6

Tabla 4.6: Muestreo doble para el control de piezas no conformes

Muestra N°	$n_1$	$d_1$	Acción <sup>(*)</sup>	$n_2$	$d_2$	$d_1 + d_2$	Acción <sup>(*)</sup>
1	50	0	A				
2	50	0	A				
3	50	1	A				
4	50	0	A				
5	50	2	TSM	242	1	3	A
6	50	1	A				
7	50	0	A				
8	50	0	A				
9	50	1	A				
10	50	0	A				
11	50	2	TSM	242	3	5	R
12	50	0	A				
13	50	1	A				
14	50	1	A				
15	50	2	TSM	242	4	6	R
16	50	3	R				
17	50	1	A				
18	50	1	A				
19	50	0	A				
20	50	1	A				
21	50	3	R				
22	50	2	TSM	242	2	4	A
23	50	1	A				
24	50	4	R				

(\*) A = aceptar, R = rechazar, TSM = tomar una segunda muestra

Veamos cuales son las diferencias con el muestreo simple. En el MS, el  $LSC = 2,5$  que implica que si en la muestra se encuentran 0, 1 ó 2 piezas no conformes, el proceso sigue operando supuestamente “bajo control”. En esta situación, la muestra N° 11 habría sido aceptada (no rechazada) cuando en realidad debería ser rechazada. Lo mismo ocurrió con la muestra N° 15. Nótese además, que en el muestreo doble la muestra N° 15 no necesita ser completada, dado que, inmediatamente cuando el conteo de las no conformidades de la segunda muestra sumados al de la primera muestra supera el  $LSC_2$ , se detiene el conteo y el proceso es rechazado.

El cálculo para obtener el porcentaje de ganancia del diseño estadístico del gráfico  $np$  con MD en relación con el diseño estadístico del gráfico  $np$  con MS para este caso en particular, está dado por:

$$\text{Porcentaje de ganancia} = \frac{ARL_{1(MS)} - ARL_{1(MD)}}{ARL_{1(MS)}} \times 100 = \frac{44,60 - 21,37}{44,60} = 52,09\%$$

Esto es, el porcentaje de ganancia en términos de tamaño muestral promedio, utilizando el gráfico con muestreo doble en lugar del gráfico con muestreo simple, es de aproximadamente un 52%.

#### 4.7 Comentario final

En este Capítulo se propone un nuevo procedimiento, el gráfico de control con muestreo doble (MD), cuyo objetivo es mejorar la capacidad para detectar cualquier condición de fuera de control mediante la observación de una segunda muestra. La eficacia de este procedimiento se basa principalmente en el hecho de que permite observar cambios pequeños o moderados en la proporción de unidades no conformes del proceso, sin aumentar el muestreo.

El muestreo doble es especialmente útil cuando se lo utiliza en procesos de alta calidad, en los cuales la proporción de unidades no conformes es reducida y el tamaño de la muestra también. En esta situación, una alternativa posible consiste en emplear el gráfico  $p$  mejorado en la primera etapa del muestreo, dado que este procedimiento ha mostrado una gran ventaja sobre el gráfico  $p$  habitual para atributos. La utilización del muestreo doble en una segunda etapa del muestreo, tiene, además, dos posibles ventajas. En primer lugar, permite reducir la cantidad total de inspección y en segundo lugar, le da al proceso una segunda oportunidad antes de tomar una decisión.

También se estableció el desempeño comparativo del gráfico  $np$  mejorado con muestreo doble respecto al mismo gráfico empleando el muestreo simple. Esta comparación reveló que el procedimiento de MD puede ser considerado más apropiado para la detección de los incrementos en la proporción de unidades no conformes en el proceso, debido a que ofrece mejor eficacia estadística (en términos de longitud promedio de corrida), sin aumentar el muestreo.

En conclusión, el gráfico  $np$  mejorado (desarrollado en el Capítulo anterior), empleado en la primera etapa del MD, se convierte en una opción interesante para el monitoreo de la proporción de unidades no conformes en los procesos de alta calidad.

Los resultados presentados en este Capítulo de la investigación, fueron publicados en el *Journal of Statistical Methodology*, como se muestra en el Anexo II.

## Capítulo 5:

### Gráficos de control basados en el conteo acumulado de unidades conformes

#### 5.1 Introducción

La competencia global y la necesidad de las empresas industriales de incrementar su participación en los mercados, las obliga a introducir tecnología en sus procesos para aumentar la productividad y disminuir costos. Tales requerimientos han llevado a muchas empresas a lograr estándares de calidad tan altos que la fracción de unidades no conformes, de algunos procesos de producción, se mide actualmente en partes por millón de unidades producidas (ppm).

Esta situación ha determinado la necesidad de prestar especial atención a los métodos de control empleados, dado que los procedimientos estadísticos tradicionales ya no deberían ser utilizados.

En este Capítulo se propone la revisión de un conjunto de propuestas para situaciones de este tipo. Ellas están referidas a la determinación de gráficos de control clasificados en la categoría de gráficos de conformidades acumuladas. Estos gráficos tienen a la distribución geométrica o a la distribución binomial negativa o a alguna de sus variantes como distribuciones de probabilidad subyacente. Se los conoce comúnmente como gráficos tipo CCC y entre ellos se destacan los gráficos CCC, CCC- $r$ , CCS y CQC.

Diferentes autores han trabajado en el desarrollo de estos procedimientos de control. Los más destacados son Calvin (1983), Goh (1987), Lucas (1989), Bourke (1991), Kaminsky, *et al.* (1992), Xie *et al.*, (2000), Chan *et al.* (2000, 2003), Nelson (1994), Wu *et al.* (2001), Kuralmani *et al.* (2002), Chen y Cheng (2010), Chen (2009) y Zhang *et al.* (2004, 2006, 2008, 2012).

El gráfico CCC (Cumulative Count of Conforming) considera el conteo acumulado de unidades conformes producidas hasta que se detecta una unidad no conforme. Este gráfico ha demostrado ser un método más adecuado que el gráfico  $p$  tradicional para el monitoreo de procesos de alta calidad.

A fin de aumentar la sensibilidad del gráfico CCC, surgió posteriormente el gráfico CCC- $r$ , que se basa en el recuento acumulado de unidades conformes producidas antes de que se observen  $r$  unidades no conformes.

Sin embargo, aunque estos gráficos han demostrado ser útiles en el seguimiento de procesos de alta calidad, ambos poseen la característica de incrementar el tiempo medio en dar una señal de

alarma cuando el proceso comienza a deteriorarse. Para evitar esta dificultad, algunos autores (Xie *et al.*, 2000; Chan y Wu, 2009, Chan, 2003, Chen y Cheng, 2010, entre otros), propusieron un cambio en el procedimiento de construcción de los límites de control que consiste en la incorporación de un coeficiente de ajuste a los límites previamente definidos. Estos límites de control ajustados se basan en la maximización de la longitud media de corrida (ARL) cuando el proceso se encuentra bajo control. En otra investigación, J.T. Chen (2009), propone una nueva aproximación para determinar los límites de control del gráfico CCC- $r$ , basado en un procedimiento de optimización a efectos de obtener un ARL cuasi insesgado y cuasi maximal.

Con la intención de establecer el procedimiento más adecuado, se realizó un estudio computacional de validación estadística, en función del ARL bajo control, a efectos de evaluar el desempeño de ambas propuestas. Los resultados obtenidos se muestran en la sección 5.3.5. Posteriormente, en la sección 5.3.6, se presenta la aplicación del gráfico CCC- $r$  a un proceso real con datos de una planta de autopartes, con análisis y discusión de los resultados.

Para la aplicación de un gráfico de control basado en el conteo de conformidades acumuladas, se necesita que las unidades producidas sean inspeccionadas una por una, es decir, secuencialmente, siguiendo el orden de producción.

Cuando la inspección se realiza por muestreo, el procedimiento consiste en controlar el número de muestras acumuladas hasta que se detecta un número determinado de productos no conformes. Esto nos lleva a un nuevo tipo gráfico de control llamado gráfico CCS (Cumulative Conforming Samples), que incluye, naturalmente, a los gráficos CCC y CCC- $r$  como casos especiales. El fundamento teórico de este gráfico se presenta en la sección 5.4, comentando sus ventajas y desventajas.

Otro gráfico, considerado como la contrapartida continua del gráfico CCC, es el gráfico CQC o gráfico de control de cantidades acumuladas (Cumulative Quantity Control Charts). Este gráfico fue propuesto por Chan *et al.* (2000), y surgió como una alternativa de los gráficos  $c$  y  $u$  de Poisson. En este caso, en lugar de considerar el número de defectos o no conformidades se cuenta el número de unidades requeridas hasta observar exactamente un defecto, donde las unidades se miden en valores continuos. Además de monitorear el proceso, este gráfico puede ser aplicado para monitorear el tiempo entre fallas (Surucu y Sazak 2009; Xie *et al.*, 2002). Este gráfico se presenta en la sección 5.5.

Cabe destacar que en el ámbito industrial, los gráficos tipo CCC, tienen múltiples aplicaciones. Por citar algunos ejemplos, se los puede emplear en procesos relacionados con el procesamiento de imágenes en los que se controlan problemas de dimensiones, formas y colores. También se los emplea en el reconocimiento de textos y lectura de códigos de barra, entre otras funcionalidades.

La implementación de los gráficos de control del tipo CCC, son de mucha utilidad en procesos que separan automáticamente de la línea de producción los productos defectuosos. Además, permiten estudiar la frecuencia de fallas y generar informes que ayuden a las decisiones de mantenimiento y al cumplimiento de normas y estándares. Otra aplicación importante es en el moldeo por inyección que se utiliza con mucha frecuencia en la construcción de piezas para el interior de automóviles, cubiertas de dispositivos electrónicos, artículos para el hogar, equipos médicos, discos compactos, etc.

Para poder analizar las propiedades de los gráficos considerados en este Capítulo, se implementaron los respectivos programas en R. De esta manera fue posible evaluar el desempeño y características de cada uno de ellos. Particularmente el programa lleva el nombre de *qcc-r* y permite construir los gráficos CCC (Xie *et al.*), CCC-*r* (Chen y Cheng y J.T.Chen) y el gráfico CCS (Zhang *et al.*). Estos programas se encuentran en el Anexo I.

En las Secciones siguientes, se presenta una revisión de los fundamentos teóricos que subyacen a cada una de las metodologías citadas, con discusión de sus ventajas y limitaciones. Además, se desarrolla la mejor propuesta para los gráficos CCC-*r*, mostrando su aplicación sobre la base de un conjunto de de un proceso industrial.

## 5.2 Gráfico CCC para inspección secuencial

El gráfico CCC fue desarrollado por Calvin (1983) para el monitoreo de procesos con muy bajo número de defectos. Se presentó y se le dio el nombre de gráfico CCC en Goh (1987). Esta es una técnica de gran relevancia para el caso del control de procesos cuando se observa un gran número de unidades conformes consecutivas entre dos no conformidades. El gráfico de control consiste en monitorear esa cantidad y las decisiones se determinan sobre la base de establecer si este número es grande o pequeño.

### 5.2.1 Fundamento teórico de los gráficos CCC

Cuando se trata con procesos cercanos a cero defectos, hay muy pocas unidades no conformes y generalmente se cuenta con una larga serie de unidades conformes antes de que una no conformidad sea observada. En situaciones como éstas, los tradicionales gráficos de control de Shewhart presentan muchas dificultades prácticas, tales como la determinación de límites de control adecuados, altas señales de falsa alarma y fracaso en la observación de mejoras del proceso.

La idea que subyace a los gráficos CCC es el hecho de que el número de unidades conformes entre dos no conformidades se modifica cuando la fracción de no conformes cambia. No es sencillo observar estos cambios en procesos cercanos a cero defectos dado que ellos suelen ser muy reducidos. Una estrategia consiste en realizar el conteo acumulado de unidades conformes entre dos

no conformidades. De esta manera, cuanto menores sean los valores absolutos graficados en el gráfico CCC, el proceso nos indicará que hay muchas unidades no conformes y por lo tanto será considerado fuera de control. Por otro lado, si la cantidad de unidades conformes entre dos no conformidades es grande, entonces el proceso probablemente ha mejorado y esto puede suceder aún en aquellos procesos con muy pocas no conformidades.

La construcción de un gráfico CCC es similar a la de un gráfico de control de Shewhart.

### 5.2.2 Límites de control

Sea  $x$  el número de unidades observadas hasta que se presente una unidad no conforme, es decir,  $(x - 1)$  unidades serán conformes y una, la  $x$ -ésima, será no conforme. Este conteo corresponde a una variable aleatoria  $X$  con distribución geométrica y a partir de ella se determinan los límites de control.

Matemáticamente, si la probabilidad de obtener una unidad no conforme es  $p$ , luego la probabilidad de obtener  $(x - 1)$  unidades conformes seguidas por una no conforme (suponiendo independencia) está dada por la siguiente función de probabilidad:

$$f(x) = p(1 - p)^{x-1}, \quad x = 1, 2, \dots \quad 5.1$$

Que corresponde a la distribución geométrica.

Posteriormente, a partir de la función de distribución, es posible obtener los límites de control considerando una probabilidad  $\alpha$  de falsa alarma. Esto es;

$$F(x) = 1 - (1 - p)^x \quad 5.2$$

Para detectar desplazamientos hacia valores bajos de  $p$  (deterioro del proceso), se necesita determinar el límite inferior de control, LIC, el cual está dado por;

$$F(\text{LIC}) = 1 - (1 - p)^{\text{LIC}} = \alpha/2 \quad 5.3$$

Luego;

$$\text{LIC} = \frac{\ln(1 - \alpha/2)}{\ln(1 - p)} \quad 5.4$$

La línea central se obtiene como:

$$F(\text{LC}) = 1 - (1 - p)^{\text{LC}} = 1/2 \quad 5.5$$

Luego;

$$LC = \frac{\ln(1/2)}{\ln(1-p)} \quad 5.6$$

El límite superior de control (mejora del proceso), LSC, está dado por;

$$F(LSC) = 1 - (1-p)^{LSC} = 1 - \alpha/2 \quad 5.7$$

Es decir;

$$LSC = \frac{\ln(\alpha/2)}{\ln(1-p)} \quad 5.8$$

El gráfico de control se construye luego con estos límites y las mediciones se grafican de la manera usual.

En el gráfico CCC se coloca el valor de  $X$  que corresponde al conteo de unidades inspeccionadas hasta que una unidad no conforme es observada. Cuando aparece una unidad no conforme el conteo se reinicia. Luego, el gráfico de control se construye graficando los conteos acumulados de unidades conformes entre dos no conformes.

### 5.2.3 Aplicación del gráfico CCC para distintos valores de $p$

El siguiente es un ejemplo obtenido por simulación para comprender como funciona el gráfico CCC para diferentes valores de  $p$ .

Supóngase un proceso en el que se inspeccionan 10.000 unidades y cada una se clasifica en una de dos posibilidades (conforme/no conforme). Además se consideran tres posibilidades para la fracción de unidades no conformes  $p$ .

1º caso:  $p = 0,01$

Tabla 5.1: Conteo de conformes entre dos no conformes del caso 1

57	55	110	15	228	25	152	30	49	85	89
39	50	16	354	182	133	36	110	50	136	59
382	26	75	111	11	181	10	19	99	42	115
81	22	58	120	63	149	37	34	20	43	7
39	84	30	61	6	55	17	300	52	12	109
7	23	11	11	189	61	343	323	48	5	150
55	64	102	158	142	34	156	38	187	767	41
70	387	13	82	113	34	115	46	54	447	
37	97	169	16	94	120	135	21	98	240	

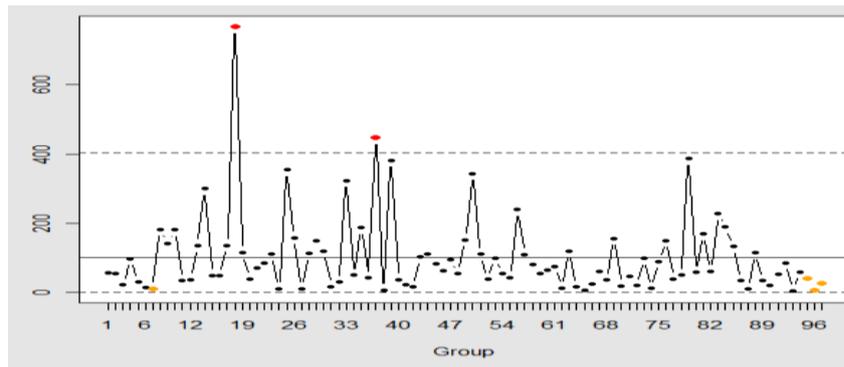


Figura 5.1: Gráfico CCC para tabla N° 5.1

2° caso:  $p = 0,001$

Tabla 5.2: Conteo de conformes entre dos no conformes del caso 2

298	1462	3170	50	301	121	407	708	1015	941	1397
-----	------	------	----	-----	-----	-----	-----	------	-----	------

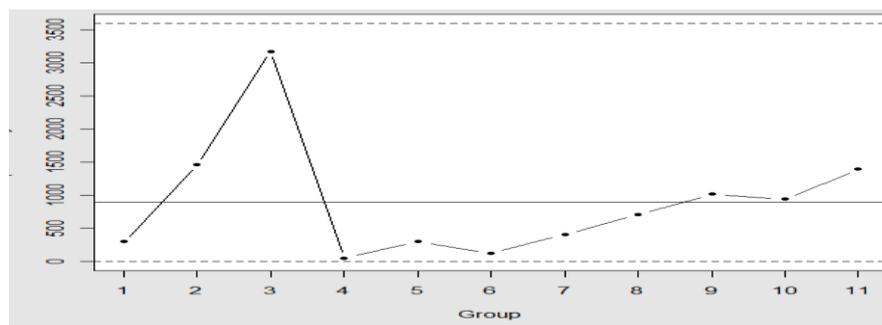


Figura 5.2: Gráfico CCC para tabla N° 5.2

3° caso:  $p = 0,0001$

Tabla 5.3: Conteo de conformes entre dos no conformes del caso 3

1083	4041
------	------

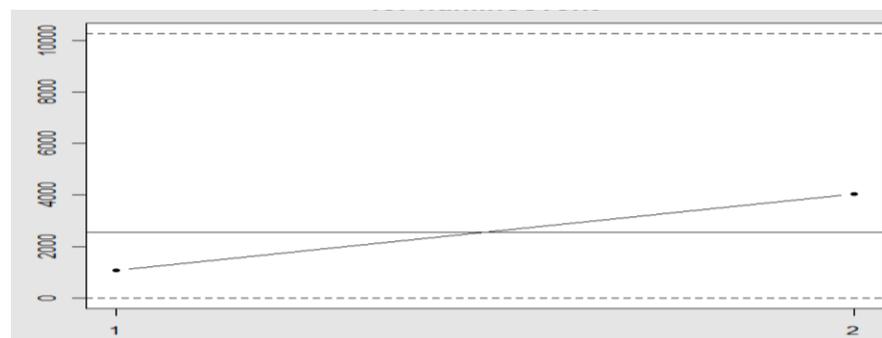


Figura 5.3: Gráfico CCC para tabla N° 5.3

Estos gráficos permiten observar que cuanto mayor es el valor de  $p$  mayor será el número de unidades no conformes y, en consecuencia, mayor será la cantidad de conteos graficados. Por este motivo, se necesitó pensar en una alternativa que no fuera tan estricta y que permitiese considerar un número mayor de no conformidades antes reiniciar el conteo. Este concepto dio origen a los gráficos CCC- $r$  ( $r \geq 2$ ) de los cuales el gráfico CCC es un caso especial, cuando  $r = 1$ .

#### 5.2.4 Decisiones basadas en el gráfico CCC

Las decisiones basadas en el gráfico CCC son similares a las decisiones tomadas con el gráfico convencional de Shewhart. La diferencia radica en que lo que se grafica es el número de unidades conformes consecutivas. Es claro que cuando el gráfico muestra una señal de fuera de control debido a un valor que supera el límite superior de control, el proceso probablemente ha mejorado.

El deterioro del proceso se determina a partir de los valores que se ubican por debajo del límite inferior de control, mostrando una cantidad reducida de conformidades entre dos no conformidades. Para la construcción del gráfico CCC se supone que  $p$ , la proporción de unidades no conformes en el proceso, se mantiene estable y que los límites de control se encuentran asociados a un cierto valor de  $\alpha$ , o probabilidad de juzgar que el proceso está fuera de control cuando no es así.

Una vez que los límites de control han sido determinados, ya no es necesario el conocimiento de los valores de  $p$  y  $\alpha$ .

#### 5.2.5 Límites de control en función de la longitud promedio de corrida

Para evaluar el comportamiento del gráfico se emplea la longitud promedio de corrida o ARL. Como fue definido en el Capítulo 2, este valor se refiere al número promedio de puntos representados en el gráfico de control antes de que sea observada una señal de fuera de control. El ARL debería ser grande cuando el proceso es considerado bajo control o cuando opera en un nivel preestablecido. Cuando este nivel se desplaza, el ARL debería disminuir de manera que la señal de alarma se determine rápidamente.

Para el gráfico geométrico con una fracción dada de no conformes  $p$ , el ARL se determina mediante la siguiente ecuación;

$$ARL = \frac{1}{1 - \beta} = \frac{1}{1 - (1 - p)^{LIC} + (1 - p)^{LSC}} \quad 5.9$$

donde  $\beta$  es el error de tipo II o la probabilidad de ausencia de alarma cuando el proceso se ha salido de control. Cuando  $p$  toma el valor  $p_0$ , que indica el nivel aceptable de calidad o el nivel medio de calidad estimado del proceso, el  $ARL = 1/\alpha$ , donde  $\alpha$  es la probabilidad de falsa alarma.

Se ha determinado (Xie, *et al.* 2002), que en los gráficos de control basados en la distribución geométrica, el conteo efectuado hasta que se produce una señal de alarma puede inicialmente incrementarse cuando el proceso se ha deteriorado (aumenta  $p$ ). En estos casos el máximo ARL no se encuentra en el valor nominal del proceso, lo cual es común en distribuciones asimétricas. El ARL es altamente sensitivo para identificar rápidos desplazamientos cuando el proceso ha mejorado o cuando se ha deteriorado. Si el proceso es estable o está bajo control, se observará una señal de falsa alarma en aproximadamente  $1/\alpha$  de las muestras o puntos representados sobre el gráfico de control. Cuando  $\alpha = 0,0027$ , que es el valor normalmente prefijado, el ARL es igual a 370,4, que debería ser constante, independientemente del promedio del proceso. Sin embargo, para gráficos geométricos, el ARL puede inicialmente incrementarse debido a la asimetría de la distribución geométrica (Xie *et al.*, 2000, 2002). Desde el punto de vista práctico, esto implica que podría tomar más tiempo para dar una señal de alarma cuando el proceso se ha deteriorado que cuando está en control.

Para dar solución a este problema, Xie *et al.* (2000, 2002) desarrollaron un procedimiento de optimización que permite obtener el máximo ARL en el valor deseado de  $p$ .

A partir de la Ecuación 5.9 el valor máximo del ARL en  $p = p_0$  se obtiene maximizando la función ARL, o lo que es lo mismo, minimizando el denominador. De esta manera, diferenciando el denominador y resolviendo para  $p$  se obtiene;

$$(1 - p_0)^{l-1} \ln(1 - p_0) \frac{dl}{dp} - (1 - p_0)^{s-1} \ln(1 - p_0) \frac{ds}{dp} = 0 \quad 5.10$$

donde  $l$  es el LIC y  $s$  es el LSC.

Luego, el valor de  $p$  que maximiza a ARL es;

$$p = 1 - \exp \left\{ \frac{\ln(1 - p_0) \ln \left[ \frac{(\alpha/2)}{(1 - \alpha/2)} \right]}{\ln \left[ \frac{\ln(1 - \alpha/2)}{\ln(\alpha/2)} \right]} \right\} \quad 5.11$$

Sustituyendo  $p$  en las Ecuaciones 5.4 y 5.8, se obtienen los nuevos límites de control modificados con el factor de ajuste, obteniéndose;

$$LIC^* = k \frac{\ln(1 - \alpha/2)}{\ln(1 - p_0)} = k LIC \quad 5.12$$

y

$$LSC^* = k \frac{\ln(\alpha/2)}{\ln(1-p_0)} = k LSC \quad 5.13$$

donde  $k$  es definido como;

$$k = \frac{\ln \left[ \frac{\ln(1-\alpha/2)}{\ln(\alpha/2)} \right]}{\ln \left[ \frac{(\alpha/2)}{(1-\alpha/2)} \right]} = \frac{\ln(LIC/LSC)}{(LSC - LIC) \ln(1-p_0)} \quad 5.14$$

De esta manera, los nuevos límites de control se obtienen multiplicando los límites originales, dados en las Ecuaciones 5.4 y 5.8, por el factor de ajuste  $k$ . Además, dado que  $k$  depende solamente de la probabilidad de falsa alarma, el promedio del proceso permanece inalterable para estos nuevos límites.

Es de destacar que nuevamente se ha establecido un procedimiento de modificación de los límites de control que no afecta a la escala original de los datos. Este procedimiento presenta características similares a las propuestas para los gráficos  $p$  y  $u$  en el Capítulo anterior, basadas en el ajuste en los límites de control mediante la aplicación de la expansión de cuantiles de Cornish-Fisher para obtener una mejor aproximación a la normal.

### 5.3 El gráfico CCC- $r$

El gráfico CCC- $r$  constituye una extensión del gráfico CCC y se utiliza para supervisar el conteo acumulado de unidades inspeccionados hasta observar  $r$  ( $\geq 2$ ) unidades no conformes. La ventaja de utilizar los gráficos CCC- $r$  han sido exploradas por muchos autores, entre los que se destacan; Kaminsky *et al.* (1992), Xie y Goh (1992,1993), Nelson (1994), Xie *et al.* (2000, 2002), Chan *et al.* (2001), Chen (2009), Chen y Cheng (2010) y Albers (2010), entre otros.

#### 5.3.1 Fundamento teórico de los gráficos CCC- $r$

El conteo acumulado de unidades conformes producidas hasta detectar  $r$  unidades no conformes constituye una cantidad que puede ser vista como la suma de  $r$  variables aleatorias independientes con distribución geométrica.

Supóngase que en un proceso de manufactura, la probabilidad de encontrar una unidad no conforme es  $p$  y sea  $X$  el conteo acumulado de unidades inspeccionadas hasta la observación de la  $r$ -ésima unidad no conforme. En esta situación, se conoce que  $X$  sigue una distribución binomial negativa con parámetros  $(r, p)$  y la función de distribución de probabilidad de  $X$  se define como;

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad x = r, r+1, \dots \quad 5.15$$

Con función de distribución acumulada de  $X$  dada por;

$$F(x, r, p) = \sum_{i=r}^x \binom{i-1}{r-1} p^r (1-p)^{i-r} \quad 5.16$$

$$= 1 - \left[ (1-p)^x + \sum_{i=1}^{r-1} \frac{1}{(i)!} p^i x(x-1) \dots (x-i+1)(1-p)^{x-i} \right]$$

Es inmediato observar que, cuando  $r = 1$ , como es el caso del gráfico CCC, la distribución se reduce al conteo distribuido geoméricamente.

Luego, cuando  $r = 1, 2$  y  $3$ , la función acumulada de  $X$  puede expresarse como;

$$F(x, 1, p) = 1 - (1-p)^x$$

$$F(x, 2, p) = 1 - [(1-p)^x + x p (1-p)^{x-1}] \quad 5.17$$

$$F(x, 3, p) = 1 - \left[ (1-p)^x + x p (1-p)^{x-1} + \frac{p^2}{2} (x^2 - x)(1-p)^{x-2} \right]$$

El gráfico CCC- $r$  tiene un mejor desempeño respecto al gráfico CCC para detectar desplazamientos en la proporción de no conformidades de un proceso. Sin embargo, a medida que  $r$  aumenta, el gráfico CCC- $r$  se torna más sensible a pequeños cambios en los valores de  $p$  especialmente cuando estos se refieren a un incremento de su valor (deterioro del proceso), (Chen y Cheng 2010).

### 5.3.2 Límites de control

Para un valor dado del error de tipo I ( $\alpha$ ), la aproximación tradicional para fijar los límites de control se basa en el criterio de probabilidades simétricas, de manera que;  $P(X < LIC|p = p_0) = P(X > LSC|p = p_0) = \alpha/2$ .

Luego, con una señal de falsa alarma predeterminada  $\alpha$  y una fracción no conforme aceptable conocida e igual a  $p_0$ , los límites de control y la línea central del gráfico CCC- $r$  pueden ser obtenidos mediante la solución de las siguientes ecuaciones;

Límite superior de control:

$$F(LSC, r, p) = \sum_{i=r}^{LSC} \binom{i-1}{r-1} p_0^r (1-p_0)^{i-r} = 1 - \alpha/2 \quad 5.18$$

Línea central:

$$F(LC, r, p) = \sum_{i=r}^{LC} \binom{i-1}{r-1} p_0^r (1-p_0)^{i-r} = 1/2 \quad 5.19$$

Límite inferior de control;

$$F(LIC, r, p) = \sum_{i=r}^{LIC} \binom{i-1}{r-1} p_0^r (1-p_0)^{i-r} = \alpha/2 \quad 5.20$$

Cuando un punto cae fuera de los límites de control, el proceso se considera, obviamente fuera de control. Sin embargo, al igual que con el gráfico CCC, cuando un punto cae por encima del límite superior del proceso, esto probablemente estará indicando que el proceso ha mejorado. Por otro lado, cuando un punto cae por debajo del límite inferior es una indicación de deterioro del proceso.

En los gráficos CCC- $r$ , la determinación del valor de  $r$  suele ser considerada como una cuestión subjetiva similar a la determinación del tamaño de muestra en los gráficos convencionales de Shewhart. Los límites de control se incrementan rápidamente cuando la fracción de no conformes  $p$  se aproxima a cero. También se incrementan a medida que  $r$  es mayor considerando un mismo valor de  $p$ . En la práctica  $r$  no debería ser muy grande dado que se necesitarían acumular más observaciones antes de que pueda tomarse una decisión. Por este motivo, se sugiere comúnmente que el valor de  $r$  no supere a 5 (Xie *et al.*, 1999, 2002).

### 5.3.3 Límites de control en función de la longitud promedio de corrida

Para evaluar el desempeño de los gráficos CCC- $r$ , también se utiliza la longitud promedio de corrida (ARL).

Para el gráfico CCC- $r$  se puede calcular el valor del ARL fuera de control ( $ARL = 1/(1 - \beta)$ ) cuando la proporción de no conformes toma un valor particular  $p$  ( $p \neq p_0$ ) como se determina en la siguiente ecuación;

$$ARL = \frac{1}{\{1 - [F(LSC|r, p) - F(LIC|r, p)]\}} \quad 5.21$$

[Como antes, cuando  $p = p_0$ , el valor del ARL es igual a  $1/\alpha$ ]

Una señal de alarma puede ser verdadera o falsa. Más precisamente, una señal de alarma es falsa cuando  $p = p_0$  y es verdadera cuando  $p \neq p_0$ . Esto hace que el ARL pueda ser categorizado

como un ARL bajo control ( $ARL_0|p = p_0$ ) y como ARL fuera de control ( $ARL_1|p \neq p_0$ ), tal como fue dado en el Capítulo anterior.

Para que un gráfico de control opere de manera adecuada debería tener valores del ARL fuera de control a lo sumo igual a los valores del ARL bajo control, para cualquier cambio en los parámetros. Un ARL con estas propiedades es usualmente llamado insesgado.

Sin embargo, existe una dificultad con las curvas ARL en los gráficos CCC- $r$  (incluso cuando  $r = 1$ ), debido a que logran su máximo en un valor de  $p > p_0$ , como puede observarse en la Figura 5.4. Esta propiedad implica que los gráficos CCC- $r$  tienen un ARL sesgado para detectar incrementos en la proporción de no conformidades, lo que implica prácticamente que podría tomar más tiempo para dar una alarma cuando el proceso se ha deteriorado.

Para comprender mejor esta situación, considérese el siguiente ejemplo obtenido por simulación. Supóngase un proceso que opera con  $p_0 = 500$  ppm. En la Tabla 5.4, se muestran los límites de control del gráfico CCC- $r$  junto con los correspondientes valores del ARL para  $p_0 = 500$  ppm y  $\alpha = 0,0027$  (los valores del ARL bajo control se han resaltado).

Tabla 5.4: Valores del ARL para gráficos CCC- $r$  con  $r = 2, 3$  y  $4$ ,  $p_0 = 500$  ppm y  $\alpha = 0,0027$

	CCC-2	CCC-3	CCC-4
$p$ (en ppm)	LIC = 107 LSC = 17797	LIC = 425 LSC = 21735	LIC = 932 LSC = 25357
100	2.1327	1.5879	1.3336
200	7.6978	5.2212	3.9208
300	32.3464	23.4420	18.0999
400	133.7300	115.2925	101.0920
<b>500</b>	<b>361.2057</b>	<b>365.9992</b>	<b>367.7335</b>
600	435.0712	395.0031	354.5421
700	357.6406	277.8730	218.5882
800	280.2157	193.6524	138.4068
900	223.5075	140.4483	92.9337
1000	182.3915	105.6440	65.5432

El ARL es altamente sensitivo para identificar rápidamente cambios cuando el proceso se ha deteriorado. En un proceso que es estable y que se encuentra bajo control, una señal de falsa alarma se producirá en aproximadamente cada 370 muestras o puntos graficados en el gráfico de control cuando  $\alpha = 0,0027$ , y esto debería ser constante independientemente del promedio del proceso. Sin embargo, en este tipo de gráficos, el ARL puede inicialmente incrementarse cuando el proceso se ha deteriorado.

Esta situación puede observarse en la Tabla 5.4, en la que se muestra un proceso centrado en 500 ppm. Como  $\alpha = 0,0027$ , el ARL debería ser igual a 370,4. Sin embargo, el máximo ocurre cuando  $p = 600$  ppm, indicando con ello que si el proceso se desplaza de 500 ppm a 600 ppm, será necesario graficar 435 puntos antes de detectar la alarma. También puede observarse que el valor del ARL para el caso de deterioro del proceso (cuando aumenta  $p$ ) es mucho mayor que el valor del ARL para un desplazamiento hacia valores menores de  $p$  (mejora del proceso).

La Figura 5.4 muestra esta área indeseable bajo la curva del ARL para el gráfico CCC-2 entre los puntos  $ARL(p_0)$  y máximo  $ARL(p_1)$ .

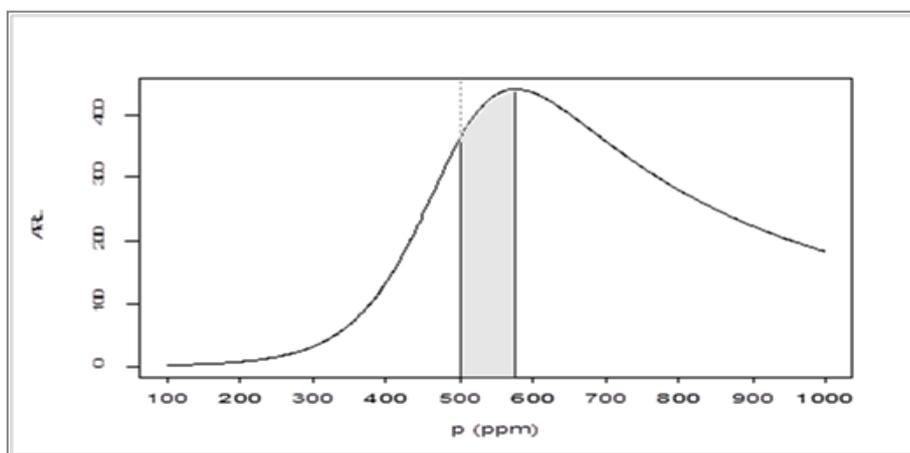


Figura 5.4: Curva ARL para el gráfico CCC-2 con máximo en  $p_1 > p_0$ .

Observando los valores del ARL para el gráfico CCC-2, es posible determinar que si el proceso se desplaza a 700 ppm, (lo que representa un deterioro del 40%), luego al menos 358 puntos o muestras deberán ser graficados para detectar ese desplazamiento. Esto es, si una unidad es inspeccionada cada minuto, tomará casi 6 horas para que el gráfico detecte la señal de fuera de control. Antes de la alarma, el proceso que está deteriorado, será considerado bajo control. Esto implica en la práctica que puede tomar más tiempo en dar la alarma cuando el proceso está deteriorado que cuando el proceso está bajo control.

Como conclusión; aunque los gráficos CCC- $r$  son más apropiados para controlar procesos de alta calidad muestran, sin embargo, un comportamiento del ARL sesgado. Esta situación tiene serias consecuencias prácticas debido a que, cuando el proceso se desplaza de su valor nominal, la salida de control no puede detectarse rápidamente.

Debido a esta complicación, han surgido en la literatura dos propuestas, la propuesta de Chen y Cheng (2010) y la propuesta de Jung Tai Chen (2009), que se pasan a considerar.

### 5.3.4 Fundamento teórico de las dos nuevas propuestas

Motivados por Xie *et al.* (2000) y Zhang *et al.* (2004), Chen and Cheng (2010) proponen una nueva aproximación a efectos de obtener los límites de control para los gráficos CCC- $r$  sobre la base de un ARL insesgado, cuando  $r \geq 2$  (que se llamará Método I). Para determinar los límites de control modificados suponen que  $p_0$  es conocido o que puede ser estimado con datos históricos del proceso en estado de control estadístico (Fase I).

Chen y Cheng (2010), desarrollaron los límites de control a partir del ARL, como se muestra en la siguiente ecuación;

$$ARL = 1 / \left\{ \begin{aligned} & 1 - \left[ (1-p)^{LIC_r} + \sum_{i=1}^{r-1} \frac{1}{(i)!} p^i LIC_r(LIC_r - 1) \dots (LIC_r - i + 1)(1-p)^{LIC_r-i} \right] \\ & + \left[ (1-p)^{LSC_r} + \sum_{i=1}^{r-1} \frac{1}{(i)!} p^i LSC_r(LSC_r - 1) \dots (LSC_r - i + 1)(1-p)^{LSC_r-i} \right] \end{aligned} \right\} \quad 5.22$$

Diferenciando la Ecuación.5.22 con respecto a  $p$ , se puede obtener el máximo ARL en  $p = p_0$ , si se encuentra un par de límites de control determinados  $LIC_r^*$  y  $LSC_r^*$  que hacen a la ecuación 5.23 aproximadamente igual a cero cuando  $p = p_0$ . Esto es;

$$\left[ \prod_{i=1}^r (LIC_r^* - i + 1) \right] (1 - p_0)^{LIC_r^*-r} - \left[ \prod_{i=1}^r (LSC_r^* - i + 1) \right] (1 - p_0)^{LSC_r^*-r} = 0 \quad 5.23$$

Por lo tanto, los límites de control óptimos,  $LSC_r^*$  y  $LIC_r^*$  pueden derivarse de los límites de control originales dados en 5.18 y 5.20, multiplicándolos por un factor de ajuste  $k_r$  mediante la siguiente relación;

$$LSC_r^* = k_r LSC_r \quad \text{y} \quad LIC_r^* = k_r LIC_r \quad 5.24$$

En esta situación, el factor de ajuste  $k_r$  no es un valor fijo como en el caso  $r = 1$  (Ecuación 5.14), sino que se obtiene mediante una ecuación de regresión especificando una determinada tasa de falsa alarma  $\alpha$  y una proporción de unidades no conformes bajo control,  $p_0$ .

Chen y Cheng (2010), publicaron varias tablas con los valores de  $k_r$  para diferentes valores de  $p_0$ . También determinaron las ecuaciones de regresión para  $k_r$  cuando  $r = 2, 3, 4$  y  $5$  en función de  $p_0$  y  $\alpha$ . Sin embargo, los autores destacan que, aunque este procedimiento produce gráficos CCC- $r$  con ARL insesgado, el ARL bajo control puede ser mayor que el valor esperado de  $1/\alpha$ . Esta situación se observa en la Tabla 5.5.

Casi en simultáneo, Jung Tai Chen (2009), propuso otra aproximación (que se llamará Método II), para determinar los límites de control del gráfico CCC- $r$ . Mediante un procedimiento de optimización, obtiene un  $ARL_0$  cuasi insesgado y cuasi maximal. Su procedimiento consiste en encontrar un intervalo  $I = [LI, LS]$  en un conjunto admisible  $S$  tal que la pendiente del  $ARL_0$  sea aproximadamente igual a cero.

Los límites de control inferior y superior (LI y LS) se obtienen a partir de la función acumulada de la distribución binomial negativa. La idea básica que subyace a su aproximación puede ser expresada por el siguiente modelo matemático;

$$\begin{aligned} &\text{Maximizar } \varepsilon = \text{la pendiente de } ARL(p_0|I) \\ &I=[LI,LS] \in S \end{aligned} \quad 5.25$$

donde LI es el límite inferior de control y LS el límite superior de control.

El modelo está sujeto a las siguientes restricciones:

$$1 - \alpha \leq P(LI \leq X \leq LS | p = p_0) < 1; \quad P(LI + 1 \leq X \leq LS | p = p_0) < 1 - \alpha \quad 5.26$$

Además;

$$\text{la pendiente de } ARL(p_0|I) \leq 0 \quad 5.27$$

Los parámetros de diseño del modelo matemático (Ecuaciones 5.25 a 5.27) son los límites de control inferior y superior. La Ecuación 5.26 asegura que el error tipo I será siempre próximo a  $\alpha$ , pero no mayor. Además, en el caso en que  $ARL(p_0|I)$  no sea máximo, la ecuación 5.27 asegura que la pendiente de  $ARL(p_0|I)$  estará tan cercana a cero como sea posible y la Ecuación 5.25 asegura que  $ARL(p_0|I)$  es casi maximal.

Respecto a la pendiente de  $ARL(p_0|I)$ ; dado un intervalo de control  $I=[LI,LS]$ , por definición  $\beta(p|I) = F(LS|r, p) - F(LI - 1|r, p)$ , y  $ARL(p|I) = 1/1 - \beta(p|I)$ . Si  $(\cdot)_p^{[n]}$  es la derivada de orden  $n$  de la función  $(\cdot)$  con respecto a  $p$ , luego, la pendiente de  $ARL(p_0|I)$  puede calcularse como;

$$\begin{aligned} ARL_p^{[1]}(p_0|I) &= \frac{1}{(1 - \beta(p_0|I))^2} \beta_p^{[1]}(p_0|I) \\ &= \frac{1}{(1 - \beta(p_0|I))^2} F_p^{[1]}(LS|r, p_0) - F_p^{[1]}(LI - 1|r, p_0) \end{aligned} \quad 5.28$$

La derivada de primer orden de  $F$  con respecto a  $p$  se obtiene a partir de la siguiente forma explícita;

$$F_p^{[1]}(x|r, p) = x \binom{x-1}{r-1} (1-p)^{x-r} p^{r-1} \quad \text{para } x = r, r+1, \dots \quad 5.29$$

Luego, a partir de la forma explícita de  $F_p^{[1]}(x|r, p)$  se puede establecer si  $ARL_p^{[1]}(p_0|I)$  es maximal.

Para la optimización, J. T. Chen propone una versión discreta del algoritmo de Newton-Raphson que describe con detalle en su artículo.

Recientemente este autor ha publicado una nueva investigación en la que propone un procedimiento para fijar los límites de control en términos del número promedio de ítems inspeccionados (ANI) en lugar del ARL.

### 5.3.5 Estudio de validación estadística de ambos procedimientos

En esta sección se realizará una validación computacional a efectos de comparar ambos procedimientos de construcción del gráfico CCC- $r$ . Por simplicidad, llamaremos a estos procedimientos con el nombre de Método I (Chen y Cheng) y Método II (J.T. Chen).

Para efectuar la comparación entre ambos métodos se tiene en cuenta que cuando el proceso se encuentra bajo control, la longitud promedio de corrida es igual a  $1/\alpha$ . Luego, si  $\alpha = 0,0027$ , el ARL debería ser igual a 370,4. A medida que  $p$  se aleja del valor de la hipótesis nula ( $p = p_0$ ), el  $ARL = 1/1 - \beta$ , es decir, la inversa de la potencia de la prueba. Por este motivo, cuanto menor es el valor de ARL mayor es la potencia y la prueba es más sensitiva.

En la Tabla 5.5 y en las Figuras 5.5 a 5.7, se muestran las curvas ARL para los gráficos CCC- $r$  con  $r = 2, 3$  y  $4$ , considerando los límites de control para un proceso con  $p_0 = 500$  ppm y  $\alpha = 0,0027$ .

Hay varias cuestiones que pueden ser analizadas a partir de la Tabla 5.5. Primero, es posible observar que, para todos los gráficos CCC- $r$ , el Método II tiene un límite inferior de control (LIC) mayor que el del Método I, pero un límite superior de control (LSC) menor. Esto implica que el Método II es más eficiente para detectar el deterioro del proceso. En la práctica, la habilidad de un gráfico de control para detectar el deterioro del proceso debería ser más importante que la habilidad para detectar su mejora.

Segundo, también se puede observar que el Método II toma valores de ARL menores que los del Método I cuando el proceso se ha deteriorado (mayor proporción de no conformes) y levemente menores cuando el proceso mejora (menor proporción de no conformes). Por lo tanto, el Método II es más sensible (mayor potencia) para reflejar el deterioro del proceso pero levemente menos sensible (menores diferencias) para reflejar la mejora del proceso.

Tabla 5.5: Curvas comparativas de ARL para los gráficos CCC- $r$  ( $\alpha = 0,0027$ )

$p$ (ppm)	CCC-2		CCC-3		CCC-4	
	Método I	Método II	Método I	Método II	Método I	Método II
	LCL = 123 UCL = 20540	LCL = 137 UCL = 20114	LCL = 470 UCL = 24055	LCL = 496 UCL = 23703	LCL = 1008 UCL = 27418	LCL = 1042 UCL = 27120
100	2.5536	2.4817	1.7599	1.7318	1.4189	1.4058
200	11.8715	11.0806	7.0648	6.7414	4.9139	4.7527
300	63.4499	56.5927	39.1392	36.0786	27.3289	25.7005
400	271.2883	228.6692	211.0947	186.7608	171.4081	156.0090
<b>500</b>	<b>444.5907</b>	<b>361.8255</b>	<b>426.0639</b>	<b>366.4419</b>	<b>413.7708</b>	<b>367.3641</b>
600	368.2718	299.0529	319.2271	274.7466	281.5628	250.4789
700	278.1593	226.1879	212.1365	183.0321	167.2752	149.2703
800	215.1696	175.1539	147.3444	127.3816	106.1691	94.9911
900	171.4403	139.6889	107.1350	92.7943	71.6774	64.2947
1000	140.0034	114.1791	80.8404	70.1496	50.8380	45.7169

Las Figuras 5.5 a 5.7, muestran las curvas de ARL para los dos Métodos. En ambos casos, las curvas logran su máximo en  $p = p_0$ , indicando que el ARL es maximal e insesgado. Algunas pequeñas desviaciones se deben a la naturaleza discreta de los límites de control de los gráficos CCC- $r$ .

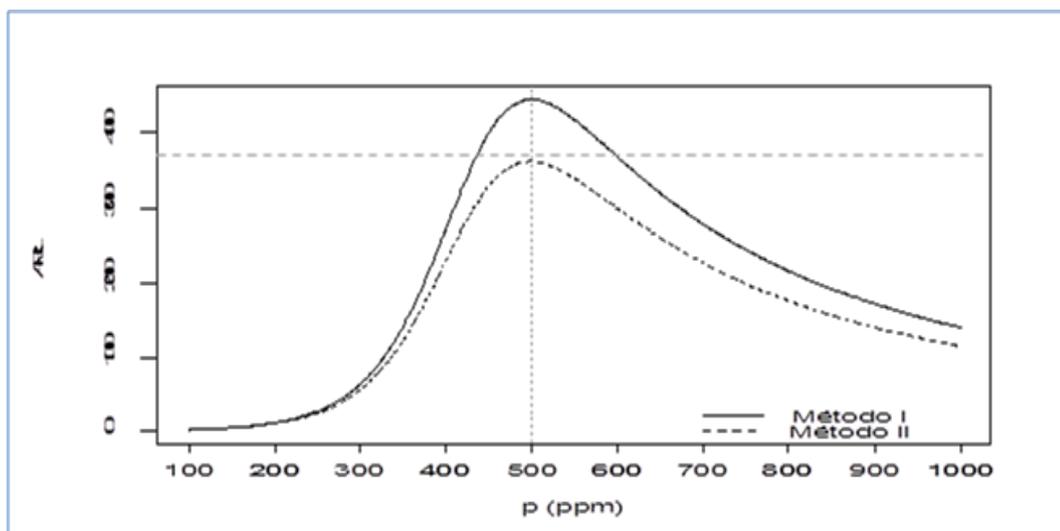


Figura 5.5: Curvas ARL para CCC-2 con ambos Métodos

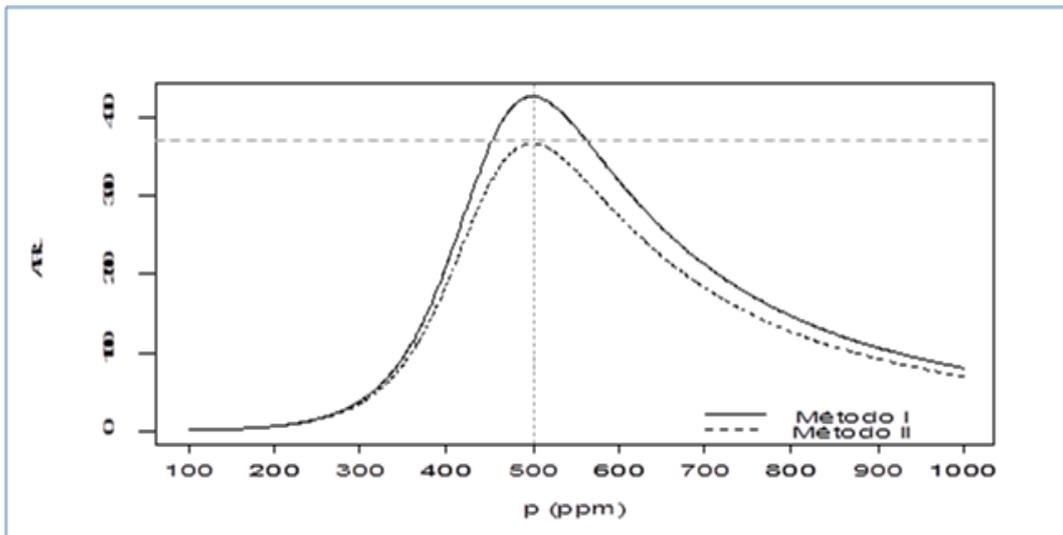


Figura 5.6: Curvas ARL para CCC-3 con ambos Métodos

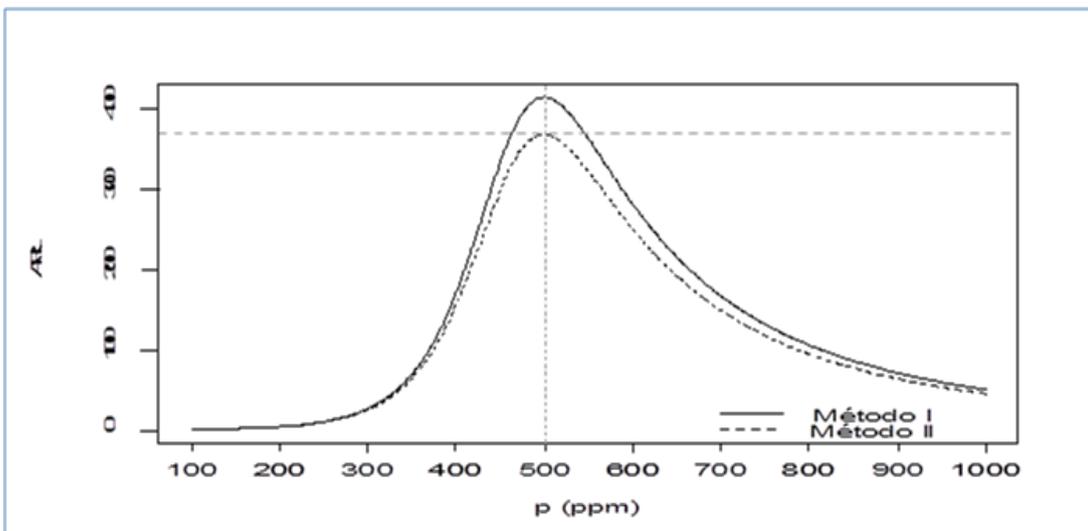


Figura 5.7: Curvas ARL para CCC-4 con ambos Métodos

### 5.3.6 Aplicación a un problema con datos de una planta de autopartes

Una empresa de autopartes produce una pieza para el panel de instrumentos de automóviles mediante un proceso de moldeo por inyección a un ritmo constante. La medición de calidad de interés es la presencia de deformaciones que impactan directamente sobre la calidad del producto. Cuando una pieza tiene alguna deformación es considerada no conforme. Una vez que el proceso de moldeo fue estabilizado, se pasó a la producción totalmente automatizada. Un lector óptico examina cada pieza que sale del molde y la descarta en presencia de alguna falla de forma. La mayoría del tiempo, el proceso permanece en estado de control, pero ocasionalmente, la fracción

de no conformes se desplaza de su nivel original a valores mayores, por lo que se ha decidido implementar el control estadístico.

Las características del proceso son las siguientes; el tiempo de ciclo de cada pieza es de 3 segundos por lo que la tasa de producción es de aproximadamente 1200 piezas por hora y de casi 10.000 piezas por día. Dada la alta tasa de producción y el reducido número de no conformidades, el Departamento de Calidad decide implementar un gráfico CCC- $r$ .

Para la determinación de los límites de control (Fase I), se tomaron 40.000 observaciones (aproximadamente 4 días de producción), obteniendo una tasa de no conformidades de 1500 ppm. Posteriormente para establecer el valor de  $r$ , se consideró que si el proceso continúa comportándose como en Fase I, en promedio, una pieza no conforme será obtenida cada 1500 piezas producidas. Es decir, cada 1500 piezas (algo más de una hora de producción) un punto deberá ser colocado en el gráfico correspondiente al conteo de piezas conformes. Como el Dpto. de Calidad considera excesivo este control, se decide graficar el conteo cada 2,5 horas de producción, por lo que se toma un  $r = 2$ . Una vez determinado el gráfico correspondiente, se tomaron 20.000 nuevas observaciones del proceso, que corresponden a dos días completos de producción, y se graficaron los conteos de conformes para ver la evolución.

Por otra parte, como a esta altura del control, es más importante para la empresa establecer si el proceso se ha deteriorado que si ha mejorado, se decide implementar el Método II por ser el procedimiento más eficiente para detectar la salida de control por mal funcionamiento del proceso.

En la Tabla 5.6 se muestran los resultados obtenidos después de una semana de control automatizado incluyendo las Fases I y II. En ella se observa el conteo de piezas conformes hasta obtener 2 no conformes. Estos resultados aparecen luego graficados como se muestra en la Figura 5.8 que corresponde al gráfico CCC-2 para el proceso de moldeo de piezas.

Tabla 5.6: Conteo de conformes hasta obtener 2 no conformes.

Datos de una semana de producción (60.000 registros)

2166	2995	1103	207	1209
1983	2761	1131	2581	1337
2528	224	1056	402	461
584	552	810	779	945
1374	1797	759	640	184
1383	439	1298	167	440
1713	342	723	276	886
498	1254	1497	450	856
1556	788	297	1047	968
1659	2076	1050	679	427
313	1274	1092	284	
885	292	1805	659	

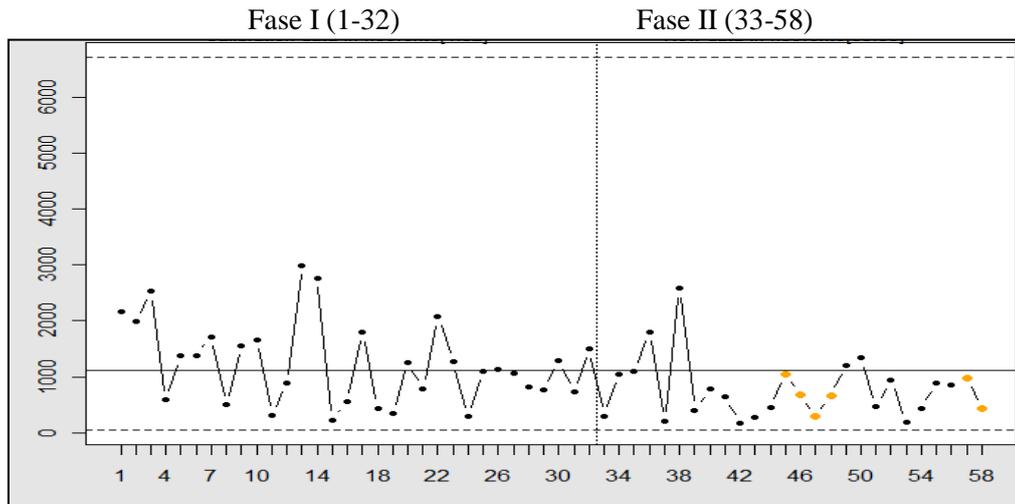


Figura 5.8: Gráfico CCC-2 para el proceso de moldeo por inyección.

Los límites de control de la Figura 5.8 son [46; 6705]. Como puede observarse el proceso se encuentra en estado de control. Sin embargo, a partir de la observación 39, prácticamente todos los conteos de unidades conformes se ubican por debajo de la línea central evidenciando un aumento en la tasa de no conformidades. Esta situación determina que el Dpto. de Calidad deba tomar acciones sobre la inyectora.

A continuación, se tratará el caso de la aplicación de los gráficos de conteo acumulado pero considerando la situación en que el proceso se monitorea no en forma secuencial (uno por uno), sino por muestreo.

#### 5.4 Gráfico CCS para inspección por muestreo

Los gráficos de control de conteo acumulado de unidades conformes, generalmente asumen implícitamente que los productos son inspeccionados de manera secuencial, unidad por unidad, siguiendo el orden de producción. Sin embargo, muchos procesos son monitoreados muestra por muestra, en cuyo caso se pierde el orden original de las unidades producidas. La inspección por muestreo es un procedimiento habitual en las empresas, principalmente cuando se trata de procesos con altos volúmenes de producción y necesidad de economizar los controles.

En este tipo de situaciones, los gráficos convencionales CCC y CCC- $r$  no se deberían emplear dado que los supuestos subyacentes de la distribución geométrica y binomial negativa ya no se mantienen.

Debido a que el orden original de los datos no se conserva cuando la inspección se realiza por muestreo, una alternativa consiste en monitorear el número acumulado de muestras hasta que se detecta un número especificado de productos no conformes.

Con la intención de dar respuesta a problemas de este tipo, en el año 2005, Zhang *et al.*, publicaron un trabajo referido a la construcción de un gráfico para monitorear el número acumulado de muestras (incluyendo la última) hasta obtener un número especificado de unidades no conformes. Este gráfico recibió el nombre de gráfico de “conteo acumulado de muestras conformes” (en inglés, Cumulative Conforming Sample o CCS), que naturalmente incluye a los gráficos CCC y CCC- $r$  como casos especiales.

Para determinar los límites de control, se considera un proceso de alta calidad con una fracción de no conformes  $p$ . Además, la inspección del proceso se efectúa mediante un procedimiento de muestreo, tomando muestras de tamaño  $n$ . Sin embargo, dentro y entre las muestras no se conserva el orden de la producción original, lo que se genera uno de los problemas del método.

Los gráficos de conteo acumulado requieren del orden de producción de manera de poder establecer un número que indique cuantas unidades conformes se han producido antes de detectar una no conforme. Cuando la inspección se realiza por muestreo, se observan  $n$  unidades por intervalo de tiempo, por ejemplo por hora de producción, por lo que muchas unidades no son inspeccionadas. Esta situación genera dudas respecto al conteo de unidades conformes entre no conformidades y es uno de los motivos por los que no ha tenido mucha aceptación en el ámbito de la calidad lo que se evidencia en la escasez de publicaciones al respecto.

El fundamento teórico de este procedimiento (Zhang *et al.*, 2005), es el siguiente; si  $M$  es el número acumulado de muestras inspeccionadas (incluyendo la última) hasta encontrar la  $r$ -ésima unidad no conforme, y  $F_M$  es la función de distribución de la variable  $M$ , luego, dada una tasa de falsa alarma  $\alpha$ , los límites de control se determinan de la siguiente manera;

$$F_M(\text{LIC} - 1) = \alpha \quad \text{y} \quad F_M(\text{LSC}) = 1 - \alpha \quad 5.30$$

Donde la función  $F_M$  surge de la siguiente expresión;

$$\begin{aligned} F_M(m) = P_r(M \leq m) &= 1 - \sum_{i=0}^{r-1} \binom{mn}{i} p^i (1-p)^{m-i} \\ &= \sum_{i=r}^{mn} \binom{mn}{i} p^i (1-p)^{m-i} \quad \text{para } r \leq mn \end{aligned} \quad 5.31$$

siendo  $p$ , la fracción de no conformes del proceso,  $r$  el número de unidades no conformes observadas y  $n$  el tamaño de la muestra.

Puede notarse que el gráfico CCS se reduce al gráfico CCC cuando  $n = r = 1$ , y al CCC- $r$  cuando  $n = 1$  y  $r > 1$ .

Cuando  $r = 1$ , la función de distribución de  $M$ , denotada por  $F_M(m/1)$  está dada por;

$$F_M(m/1) = 1 - (1 - p)^m \quad 5.32$$

A partir de esta ecuación es posible obtener el límite inferior de control de la siguiente manera;

$$\begin{aligned} F_M(\text{LIC} - 1) &= 1 - (1 - p)^{n(\text{LIC}-1)} = \alpha \\ n(\text{LIC} - 1) \ln(1 - p) &= \ln(1 - \alpha) \\ \text{LIC} &= \frac{\ln(1 - \alpha)}{n \ln(1 - p)} + 1 \end{aligned} \quad 5.33$$

mientras que el límite superior de control se obtiene como;

$$\begin{aligned} F_M(\text{LSC}) &= 1 - (1 - p)^{n(\text{LSC})} = 1 - \alpha \\ n(\text{LSC}) \ln(1 - p) &= \ln(\alpha) \\ \text{LSC} &= \frac{\ln(\alpha)}{n \ln(1 - p)} \end{aligned} \quad 5.34$$

Cuando  $r = 1$ , no es dificultoso encontrar una solución para el cálculo de los límites de control del gráfico CCS. Sin embargo, cuando  $r > 1$ , el cálculo se complica y se necesitará recurrir a algún programa de computación para obtener una solución aproximada.

Es fácil ver la similitud de estos límites con los dados en las ecuaciones 5.4 y 5.8.

Al igual que como fue establecido para los gráficos CCC y CCC- $r$ , en este caso también se observó que el ARL es sesgado. Por este motivo, en un trabajo reciente, Zhang *et al.* (2012) proponen un procedimiento de control mejorado para procesos de alta calidad bajo grupos de inspección, cambiando el nombre del gráfico de CCS a CCC<sub>G</sub> (Cumulative Count of Conforming Groups). Ellos también establecieron un procedimiento para determinar los límites de control a partir de la maximización de la longitud media de corrida (ARL) cuando el proceso se encuentra bajo control.

La longitud promedio de corrida ARL se define, en este caso, como el número de puntos de la estadística  $M$  que deben ser graficados antes de que un punto indique una señal de fuera de control. Luego;

$$ARL = \frac{1}{1 - (1 - p)^n_{LIC} + (1 - p)^n_{(LSC-1)}} \quad 5.35$$

Para mejorar la performance del ARL, Zhang *et al.* (2012), desarrollaron un procedimiento de corrección del sesgo mediante una técnica similar a la empleada por Xie *et al.* (2002), como fue dada en la Sección 5.2.5.

Los nuevos límites de control están ahora dados por;

$$LIC^* = \gamma_{\alpha^*} \frac{\ln(1 - \alpha^*/2)}{n \ln(1 - p_0)} + 1 \quad y \quad LSC^* = \gamma_{\alpha^*} \frac{\ln(\alpha^*/2)}{n \ln(1 - p_0)} \quad 5.36$$

donde;

$$\gamma_{\alpha^*} = \frac{\ln \left[ \frac{\ln(1 - \alpha^*/2)}{\ln(\alpha^*/2)} \right]}{\ln \left[ \frac{(\alpha^*/2)}{(1 - \alpha^*/2)} \right]} \quad 5.37$$

A partir de esta solución, definen un nuevo valor para el ARL llamado ahora ARL<sub>u</sub>, que toma la siguiente expresión:

$$ARL_u = \frac{1}{1 - (1 - p)^{\gamma_{\alpha^*} \frac{\ln(1 - \alpha^*/2)}{\ln(1 - p_0)}} + (1 - p)^{\gamma_{\alpha^*} \frac{\ln(\alpha^*/2)}{\ln(1 - p_0)}}} \quad 5.38$$

Luego, dados  $p_0$  y  $\alpha^*$ , el gráfico CCC<sub>G</sub> con límites como en 5.38, tiene valores del ARL<sub>u</sub> insesgado, que logra su máximo en  $p = p_0$ . Sin embargo, el problema radica en la determinación del valor  $\alpha^*$ .

Zhang *et al.* (2012), establecieron que, dado un valor de  $\alpha$  (o equivalentemente un valor del ARL<sub>u</sub> bajo control), el correspondiente valor del  $\alpha^*$  puede determinarse resolviendo la siguiente ecuación;

$$\alpha = 1 - (1 - p_0)^{\gamma_{\alpha^*} \frac{\ln(1 - \alpha^*/2)}{\ln(1 - p_0)}} + (1 - p_0)^{\gamma_{\alpha^*} \frac{\ln(\alpha^*/2)}{\ln(1 - p_0)}} \quad 5.39$$

Posteriormente confeccionaron una tabla que relaciona los valores de  $\alpha^*$  y  $\gamma_{\alpha^*}$ , para valores dados de  $\alpha$  y diferentes valores de  $p_0$ .

Como puede observarse, este procedimiento se transforma en una metodología sumamente compleja para aplicar en la práctica aún empleando algún programa de computación.

Finalmente se presenta el gráfico CQC que se utiliza en situaciones en que la cantidad producida se mide en una escala numérica más que como un conteo. Este gráfico se conoce como la contrapartida continua del gráfico CCC.

## 5.5 Gráfico CQC para cantidades acumuladas

Tradicionalmente los gráficos  $c$  y  $u$  de Shewhart basados en la distribución Poisson, son utilizados para monitorear la cantidad de no conformidades o el promedio de no conformidades por unidad inspeccionada. Por ejemplo, una óptica de un automóvil puede tener 3 imperfecciones.

Sin embargo, hay otros procesos en los que esta cantidad se mide en unidades de medida continuas. Este es el caso, por ejemplo, de procesos tales como la producción de chapas de acero cuya cantidad es medida en metros cuadrados, los líquidos en litros, el alambre en metros, etc. En situaciones como estas, los gráficos  $c$  y  $u$  no deberían utilizarse cuando la tasa de no conformidades es muy baja. Esto hace necesario, por lo tanto, determinar un procedimiento alternativo para un ambiente de alta calidad. Surgió así el gráfico CQC.

El gráfico CQC (Cumulative Quantity Control Charts) opera de la siguiente manera. Supóngase que la ocurrencia de no conformidades en el proceso puede ser modelada mediante una distribución de Poisson con media  $\lambda$  que corresponde al número de no conformidades por cantidad unitaria de producto. Luego, el número de unidades  $Q$  (que no es necesariamente un entero) requeridas para observar exactamente una no conformidad es una variable aleatoria exponencial con función de probabilidad acumulada dada por;

$$F(q) = 1 - e^{-\lambda q} \quad q > 0 \quad 5.40$$

donde el valor de  $\lambda$  se puede obtener como promedio de los valores observados de  $Q$  o por el historial previo del proceso.

Nótese que  $Q$  es una cantidad asociada a una escala numérica continua. Por ejemplo, puede referirse a la cantidad de metros de tela fabricados antes de encontrar una falla.

El gráfico CQC no necesita ni del tamaño de la muestra ni del intervalo de muestreo, por lo que en su construcción sólo se deben establecer los límites de control. Como el gráfico está basado en la distribución exponencial, que es altamente asimétrica, algunos autores (Chan *et al.*, 2000, 2002; Xie *et al.*, 2002) sugieren el uso de límites probabilísticos de control en lugar de los tradicionales límites  $k$ -sigma.

Los límites de probabilidad se obtienen de tal manera que el proceso es considerado fuera de control si el número de unidades  $Q$  inspeccionadas hasta observar una no conformidad es menor

que el límite inferior de control (LIC) o mayor que el límite superior de control (LSC). Luego, dada una tasa de falsa alarma  $\alpha$  aceptable, los límites de control se determinan como:

$$\text{LIC} = -\lambda^{-1} \ln(1 - \alpha/2) \quad \text{y} \quad \text{LSC} = -\lambda^{-1} \ln(\alpha/2) \quad 5.41$$

donde la línea central está definida por el percentil del 50%, de la siguiente manera;

$$\text{LC} = -\lambda^{-1} \ln(1/2) \approx 0,6931 \lambda^{-1} \quad 5.42$$

En este gráfico, cuando  $q < \text{LIC}$ , ello indica que probablemente el proceso se ha deteriorado puesto que la cantidad de unidades inspeccionadas hasta que se observa una unidad no conforme es menor a lo tolerado. El caso contrario ocurre cuando  $q > \text{LSC}$ , puesto que en esta situación se ha demorado más en detectar una unidad no conforme de lo tolerado, indicando una probable mejora en el proceso. Del mismo modo que en el gráfico CCC, en el gráfico CQC la cantidad acumulada se reinicia cuando una no conformidad es observada.

## 5.6 Comentario final

Los gráficos de control para procesos de alta calidad, basados en el conteo acumulado de unidades conformes, han atraído recientemente mucha atención. Estos gráficos constituyen el procedimiento apropiado para tratar con procesos secuenciales (uno por uno), en los que es posible establecer el orden de producción de cada unidad.

En particular, los gráficos CCC y CCC- $r$  mostraron tener mejor desempeño que los gráficos  $p$  y  $np$  para monitorear la fracción de unidades no conformes en procesos de alta calidad. Sin embargo, la asimetría de la distribución geométrica y binomial negativa subyacentes a estos gráficos hacen que el ARL tenga un comportamiento sesgado.

Debido a esta limitación, han aparecido en la literatura dos nuevos enfoques que, mediante diferentes procedimientos, obtienen un ARL maximal e insesgado. En este Capítulo se analizaron y compararon los resultados numéricos de ambos métodos a efectos de determinar la mejor opción para la aplicación de los gráficos CCC- $r$ . Mediante un estudio de simulación se compararon las curvas ARL para los gráficos CCC- $r$  con  $r = 2, 3$  y  $4$ , considerando los límites de control para un proceso con  $p_0$  y  $\alpha$  dados. Este estudio comparativo reveló que el Método II (J.T. Chen) puede considerarse más apropiado (mayor potencia) para la detección de un aumento en la fracción de no conformes del proceso. También se mostró la aplicación con datos de una planta de autopartes.

Estos resultados dieron origen a una publicación enviada al *Journal of Process Control*, bajo el título: “Study of the CCC- $r$  Control Chart for High-Quality Processes and its Application to Real Data from a Car Spare Parts Plant”, como puede verse en el Anexo II.

Otro procedimiento considerado en este Capítulo corresponde a los gráficos CCS. Este tipo de gráficos está relacionado con aquellos procesos de alta calidad en el que las unidades producidas son inspeccionadas muestra por muestra, cuando la probabilidad de que una muestra contenga más de un producto no conforme es muy pequeña. El problema de estos gráficos está relacionado con el hecho de que, tanto dentro como entre muestras el orden original de producción se pierde. Por este motivo la estadística que monitorea el conteo de conformes entre no conformes pierde veracidad.

Al igual que los gráficos CCC y CCC- $r$ , Zhang *et al.*, (2012) establecieron que este gráfico también presenta problemas de sesgo del ARL, por lo que se desarrolló una versión mejorada que se publicó recientemente sólo para el caso  $r = 1$ . Sin embargo, se mostró en este Capítulo la complejidad del procedimiento, especialmente a la hora de determinar el valor adecuado de la probabilidad de falsa alarma  $\alpha$ .

Por último se presentó el gráfico CQC que es también una alternativa del gráfico CCC pero que se utiliza en casos en que las unidades de producción son continuas, como metros de tela, litros de aceite, etc.

Finalmente, a efectos de facilitar la utilización de estos gráficos, los mismos han sido desarrollados en lenguaje R como una extensión de los programas básicos de control de calidad que se encuentran en la librería “qcc”. Los interesados pueden encontrar el algoritmo en el Anexo III de esta investigación.

### 5.6.1 Comentario sobre otros estudios relacionados

En este Capítulo se han determinado los procedimientos generales de implementación de los gráficos CCC. Sin embargo, también existen otros estudios relacionados con este tipo de gráficos que cubren cuatro aspectos no menos importantes. El primer aspecto es el estudio de la correlación serial en los gráficos CCC. Cuando el proceso de producción muestra una correlación serial, la tasa de falsa alarma será mucho más alta que la calculada bajo el supuesto de independencia. Lai *et al.* (1998, 2000) usaron el modelo de correlación binomial de Madsen (1993) para resolver el problema de correlación del gráfico CCC y posteriormente aplicaron el modelo de Markov. Tang y Cheong (2006) proponen un nuevo gráfico para resolver el problema de correlación dentro de cada grupo inspeccionado, llamado gráfico de conteo acumulado de conformidades en grupo ( $C^4$ ). Ellos definen al grupo como un conjunto de unidades que contienen cero o una no conformidad cuando no se observaron no conformidades en los grupos anteriores. En el gráfico  $C^4$  se coloca el número

de grupos conformes antes que un grupo no conforme se registre. Chen y Cheng (2007), estudiaron el seguimiento de procesos con dependencia de Markov, aplicando a los gráficos CCC un procedimiento de muestreo variable.

El segundo aspecto es el estudio del intervalo de muestreo variable (VSI=Variable Sample Interval) para el gráfico CCC. El VSI ha sido estudiado ampliamente por muchos investigadores (Lee y Bai, 2000; Bai y Lee, 2002; Carot *et al.*, 2002; Wu y Luo, 2004; Chen, 2004; Lin y Chou, 2005; Villalobos *et al.* 2005). La razón para la aplicación del VSI es que puede reducir el costo del muestreo y aumentar la sensibilidad del gráfico con los parámetros de diseño apropiados. Teniendo en cuenta la aplicación exitosa del VSI en otros gráficos, Liu *et al.* (2006a) aplicaron el esquema de VSI al gráfico CCC.

El tercer aspecto es el estudio del comportamiento aleatorio de los desplazamientos del gráfico CCC. El desplazamiento aleatorio significa que la ocurrencia del cambio del proceso puede darse en cualquier punto entre dos no conformidades. En cambio, el desplazamiento fijo significa que la ocurrencia del cambio del proceso sólo puede darse inmediatamente después de una unidad no conforme. Wu y Spedding (1999), estudiaron las propiedades del ATS (Average time to signal) para desplazamientos fijos y aleatorios.

El cuarto aspecto es el diseño económico del gráfico CCC. Duncan (1956) fue pionero en el diseño económico de los gráficos de control. Más tarde, otros autores aplicaron un modelo más general para obtener la solución óptima. Zhang *et al.* (2008) desarrollaron un modelo económico del gráfico CCC para el caso en que el orden de producción original no se conserva. En lugar de monitorear el conteo acumulado de unidades conformes, el gráfico de Zhang monitorea el número acumulado de muestras conformes.

Respecto a los gráficos CCC- $r$ , varios autores han realizado investigaciones interesantes. Por ejemplo, Ohta *et al.* (2001), desarrollaron un modelo económico para el gráfico CCC- $r$  y propusieron un método simplificado para determinar las variables de diseño óptimo. Wu *et al.* (2001) y Zhang *et al.* (2007), estudiaron las propiedades del tiempo medio hasta la señal (ATS) para el gráfico CCC- $r$  con desplazamientos aleatorios. Para simplificar el problema de la gran cantidad de unidades inspeccionadas antes de que se pueda tomar una decisión, Chan *et al.* (2003) proponen el gráfico CCC $_{1+\gamma}$  que fue inspirado en el muestreo doble. Un gráfico CCC $_{1+\gamma}$  es un gráfico CCC en dos etapas, con una probabilidad de falsa alarma de  $(1-\gamma)\alpha$  para la primera etapa y  $\gamma\alpha$  para la segunda etapa.

También existen extensiones del gráfico CQC. Una de ellas es el gráfico CQC- $r$  que se utiliza para monitorear el número de unidades inspeccionada hasta la ocurrencia de  $r$  no conformidades. El gráfico CQC- $r$  fue propuesto por Xie *et al.* (2002b), quienes le dieron el nombre de gráfico  $T_r$ .

En este caso, en lugar de contar el número de unidades conformes hasta obtener  $r$  no conformidades, se registra el tiempo acumulado hasta observar  $r$  defectos. También existe un procedimiento en dos etapas para el gráfico CQC- $r$ .

Una limitación de los gráficos CCC es que la decisión se basa en cada conteo. Sin embargo, hay muchas formas de incorporar información previa del proceso. Un procedimiento simple es el gráfico CCC- $r$ , pero también existen procedimientos condicionales y procedimientos basados en corridas con reglas (run rules). Otra alternativa es el uso de los gráficos CUSUM (Cumulative Sum) y EWMA (Exponencial Weighted Moving Average), que aunque fueron originalmente desarrollados para datos continuos, varios autores han investigado sus propiedades para datos de atributos, entre los que se encuentran los gráficos CCC (Lucas, 1985; Chan y Gan (2001); Yeh *et al.* (2008). Sin embargo, varios estudios han mostrado que cuando se trata de procesos de alta calidad en los cuales el número de no conformidades es muy reducido, estos gráficos muestran señales de alarma en cualquier muestra que contenga una unidad no conforme dado que esencialmente son una extensión de los tradicionales gráficos  $c$  y  $np$  de Shewhart (Lucas, 1985).

## Capítulo 6:

### Gráfico de control para procesos de atributos con exceso de ceros

#### 6.1 Introducción

Debido al avance tecnológico y a la automatización de los procesos de producción, un proceso bien diseñado podría tener un número mayor de conteo de ceros de los que se esperaría encontrar bajo distribución Poisson o binomial. En particular, en el ámbito de la producción, es cada vez más frecuente encontrar muestras conformes, es decir, muestras sin ninguna unidad defectuosa. Como ya se ha mencionado en los Capítulos anteriores, estos procesos se conocen actualmente como procesos de alta calidad en los que, gracias al control continuo, se han logrado estándares de calidad realmente muy elevados.

Tradicionalmente estos procesos han sido monitoreados empleando las distribuciones Poisson o binomial, pero ambas distribuciones tienden a subestimar la media y la variabilidad del proceso en presencia de gran cantidad de ceros. Los gráficos de control resultantes tienen límites de control muy estrictos que conducen a una mayor tasa de falsas alarmas en la detección de señales de fuera de control.

Una forma de solucionar este inconveniente consiste en extender las distribuciones básicas de Poisson y binomial a una que pueda ser utilizada para modelar el efecto de una mayor cantidad de ceros. De esta manera, surgieron los procesos “zero-inflated” o procesos con “excesos de ceros”, que son aquellos procesos en los que se observa que los datos muestrales presentan una frecuencia más elevada de unidades conformes (cero defectos) que la que se debería esperar si la muestra hubiera sido generada mediante una distribución Poisson o binomial.

En los últimos años, varios investigadores han proporcionado nuevos modelos que se obtienen mediante la modificación de los modelos tradicionales existentes. Particularmente, los modelos con exceso de ceros, que incluyen a los modelos zero-inflated Poisson (ZIP) y zero-inflated binomial (ZIB), son modelos que se obtienen combinando adecuadamente el modelo tradicional con una singular distribución de ceros.

Xie *et al.*, (2001) desarrollaron un gráfico  $c$  para el modelo ZIP que denominaron gráfico  $c_{ZIB}$ . Estimaron los parámetros por máxima verosimilitud (MLE) y aplicaron varios test para probar la hipótesis de la existencia de un parámetro extra en el modelo Poisson asociado con el exceso de

ceros. Además, estudiaron la eficiencia de este gráfico de control para detectar desplazamientos del valor medio del número de no conformidades cuando el proceso se encuentra bajo control.

Sim y Lim (2008) desarrollaron varias propuestas para la construcción de gráficos de control para datos con exceso de ceros, tanto para el caso Poisson como para el binomial. Ellos recomendaron construir los gráficos de control sustituyendo los límites de control  $3\sigma$  con límites construidos sobre la base de intervalos de confianza. Para el caso Poisson, desarrollaron el gráfico  $c_j$ . El parámetro  $\lambda$ , estimado a partir del modelo con exceso de ceros, ZIP, es utilizado para construir un gráfico unilateral con el límite superior de control obtenido sobre la base del intervalo de Jeffreys. Además compararon este gráfico  $c_j$  con el gráfico  $c$  tradicional y el gráfico  $c_{ZIP}$ . En el caso binomial, el parámetro  $p$  estimado a partir de un modelo ZIB se utilizó para construir el gráfico  $np_j$  con el límite superior de control basado en el intervalo de Jeffreys, y el gráfico  $np_{BS}$  construido con el intervalo de Blyth-Still. Compararon luego estos gráficos con el gráfico  $np$  de Shewhart y con el gráfico  $np$  basado en el modelo ZIB (gráfico  $np_{ZIB}$ ). La eficiencia de los distintos gráficos se analizó empelando la longitud promedio de corrida (ARL). Sin embargo los autores advierten que los gráficos resultantes muestran un valor del ARL mucho menor que el valor deseado (nominal) debido a una deficiente cobertura probabilística de los límites de control. Además establecen que en aplicaciones biológicas (que son las que ellos presentan) los parámetro  $p$  y  $\lambda$  son generalmente grandes, mientras que en los procesos de alta calidad, son extremadamente pequeños. Por este motivo sugieren para estos casos la búsqueda de nuevos métodos.

Fatai *et al.* (2010), investigaron un gráfico de control basado en la distribución ZIB truncada (TZIB), aplicando límites de probabilidad en lugar de los límites de control de Shewhart. Para evaluar la aptitud del grafico propuesto también utilizaron el ARL calculado sólo con los valores positivos de la variable binomial. Los resultados mostraron la importancia de utilizar la distribución ZIB truncada respecto de emplear el gráfico  $np$  tradicional.

Peerajit y Mayuresawan (2010), llevaron a cabo un extenso estudio de la distribución ZIP, comparando la eficiencia de los gráficos  $c$ ,  $c_{ZIP}$  y  $c_j$  en procesos con diferente proporciones de exceso de ceros. Para la comparación emplearon el ARL y el ACP (Average Coverage Probability) como indicadores de eficiencia. Los resultados revelaron que el gráfico  $c_j$  tiene una eficacia más alta que los otros gráficos de control cuando el proceso está bajo control.

Kateme y Mayuresawan (2012a), propusieron gráficos basados en la media, mediana y varianza de la distribución geométrica, denominados, gráficos  $c_g$ ,  $c_{mg}$  y  $c_{me}$ . En el grafico  $c_g$ , los límites de control se construyeron con la media y la varianza de la distribución geométrica. En los gráficos  $c_{mg}$ , la media de la distribución geométrica se utilizó para sustituir tanto la media como la varianza en los límites de control. En el grafico  $c_{me}$ , se utilizó la mediana y la varianza de la

distribución geométrica para construir los límites de control. Los resultados de las comparaciones se presentaron en forma de Tabla en la que se establece el gráfico de control adecuado para un amplio rango de valores de los parámetros del modelo ZIP.

En otra investigación, Katemee y Mayureesawan (2012b), propusieron tres nuevos gráficos para procesos basados en la distribución ZIP. La distribución ZIP se aproximó mediante una distribución chi-cuadrada no central con parámetro  $\lambda$  ( $\lambda Chi$ ). El mejor valor de ajuste de  $\lambda Chi$  se utilizó para sustituir la media y la varianza estimada en los límites de control de los gráficos  $c$  de Shewhart empleando tres propuestas diferentes a) el gráfico  $c_{Chi}$ , en el que los valores estimados de la media y la varianza de los gráficos  $c$  se sustituyeron por  $\lambda Chi$ , b) el gráfico  $c_{CCChi}$ , donde los valores de la media y la varianza son reemplazados con los estimadores de la media y la varianza de la distribución chi-cuadrada no central, y c) el gráfico  $c_{MChi}$ , en el que la media se reemplaza por la media estimada de la distribución chi-cuadrada no central y la varianza por el rango intercuartílico. Numerosas simulaciones se llevaron a cabo para evaluar la aptitud de los nuevos gráficos empleando el ARL y el ACP. Los gráficos propuestos se compararon con los gráficos  $c$ ,  $c_J$  y  $c_{ZIP}$  y se indicó la eficiencia de cada uno de ellos cuando los procesos se encontraban bajo control y fuera de control.

Yawsaeng y Mayureesawan (2012), estudiaron la eficiencia de los gráficos de control para modelos ZIB. Propusieron varios gráficos basados en diferentes intervalos de confianza binomiales, tales como: los gráficos  $np_{ZIB}$ , los gráficos  $np_J$  con intervalo de Jeffreys, los gráficos  $np_w$  con intervalos de Wilson, gráficos  $np_{AC}$  con intervalo de Agresti-Coull y los gráficos  $np_{BS}$  con intervalo de Blyth-Still. Todos fueron comparados con el tradicional gráfico  $np$  de Shewhart. La eficiencia de los gráficos de control se evaluó en términos del ARL y también en términos de probabilidad promedio de cobertura (ACP). Como conclusión efectuaron recomendaciones respecto al gráfico de control adecuado teniendo en cuenta diferentes valores para la proporción de ceros ( $\phi$ ), distintos niveles de desplazamiento en la proporción de no conformidades ( $\theta$ ) y diferentes valores para la varianza ( $\sigma^2$ ).

Sin embargo, los resultados obtenidos de estas investigaciones son muy variables debido a que dependen de la magnitud de la proporción de ceros, de los cambios en la media o en la proporción de no conformidades, del tamaño de las muestras, de la variabilidad y del estado en el que se encuentra el proceso ya sea bajo control o fuera de control. Esta situación conduce a una cierta complejidad a la hora de seleccionar el procedimiento de control adecuado.

En este Capítulo se desarrolla una metodología más simple para determinar si un proceso es binomial con un número elevado de ceros debido a una tasa de no conformidades muy chica, o si se tiene una masa de probabilidad asociada con la presencia de ceros que es superior a la esperada

para una distribución binomial estándar. Teniendo en cuenta que la eficiencia de un gráfico de control está determinada por la bondad de ajuste entre los datos observados y la distribución asumida para la construcción del gráfico, determinar la distribución adecuada es de vital importancia para su implementación y para que éste funcione adecuadamente. Cuando se fracasa en tener en cuenta el exceso de ceros, se genera un modelo mal especificado que puede resultar en estimadores sesgados o inconsistentes.

A efectos de realizar la elección adecuada del modelo, se propone utilizar el modelo lineal generalizado por ser éste un procedimiento muy flexible para tratar con datos de conteos. El modelo lineal generalizado es una extensión del modelo lineal clásico a una familia más general que involucran una variedad de distribuciones seleccionadas de una familia exponencial, como lo son las distribuciones, normal, binomial, Poisson, Gamma y binomial negativa. Fueron introducidos por primera vez por Nelder y Wedderburn (1972) y progresivamente se han aplicado en diversas disciplinas como agricultura, ecología, economía, ingeniería y medicina, entre otras.

A continuación se presentan los modelos ZIP y ZIB. Posteriormente se definen las características del modelo lineal generalizado y su aplicación al caso binomial y binomial con exceso de ceros por ser este el menos tratado en la literatura, especialmente para el control estadístico de la calidad. Luego se define el procedimiento para seleccionar el modelo más adecuado, la construcción de los límites del gráfico de control y el estudio de la sensibilidad del gráfico.

## 6.2 Modelos con exceso de ceros o modelos cero-inflados

En una situación ideal, sólo unidades perfectas serían observadas es decir, unidades o muestras de unidades todas conformes. No obstante, una unidad o una muestra pueden estar afectadas por un mecanismo aleatorio que produce no conformidades. Si tal mecanismo, que puede ser controlable o no, ocurre, aparecerán entonces algunas no conformidades.

En el contexto del control estadístico de procesos esta situación indica que, además de la posibilidad de que el número de no conformidades en la muestra tome un valor cero bajo variación aleatoria, puede observarse un número adicional de ceros con probabilidad  $\phi$ .

La distribución del número de no conformidades en la muestra puede ser tratada como una distribución Poisson modificada, de la siguiente manera:

$$P(\text{cero no conformidades en la muestra}) = \phi + (1 - \phi) e^{-\lambda} \quad 6.1$$

$$P(y \text{ no conformidades en la muestra}) = (1 - \phi) \frac{e^{-\lambda} \lambda^y}{y!} \quad 6.2$$

Este modelo es también conocido como modelo Poisson con exceso de ceros (ZIP) y tiene media y varianza dadas por;

$$E(Y) = (1 - \phi)\lambda \quad ; \quad \text{Var}(Y) = (1 - \phi)\lambda [1 + \lambda \phi] \quad 6.3$$

Para el caso binomial, la distribución modificada se obtiene como;

$$P(\text{cero no conformidades en la muestra}) = \phi + (1 - \phi)(1 - p)^n \quad 6.4$$

$$P(y \text{ no conformidades en la muestra}) = (1 - \phi) \binom{n}{y} p^y (1 - p)^{n-y} \quad 6.5$$

Este modelo se conoce con el nombre de modelo binomial con exceso de ceros (ZIB) y tiene media y varianza dadas por;

$$E(Y) = (1 - \phi)np \quad ; \quad \text{Var}(Y) = (1 - \phi)np [(1 - p) + np\phi] \quad 6.6$$

Es claro que cuando  $\phi$  tiende a 0, las distribuciones ZIP y ZIB toman sus distribuciones originales.

Sim y Lim (2008), desarrollaron los modelos con exceso de ceros a partir de la definición de una variable aleatoria cero-inflada  $Y$  que puede expresarse como;

$$Y = \begin{cases} 0, & \text{con probabilidad } \phi \\ X, & \text{con probabilidad } (1 - \phi) \end{cases} \quad 6.7$$

donde  $0 \leq \phi \leq 1$  y  $X$  es la variable aleatoria discreta de interés con función de probabilidad  $g(x; \theta) = P(X = x)$ , donde  $\theta$  es el vector de parámetros. Es decir, además de la posibilidad de que la variable aleatoria  $X$  tome un valor cero bajo variación aleatoria, un número adicional de ceros puede ocurrir con probabilidad  $\phi$ . La función de probabilidad de la variable  $Y$  está dada por;

$$P(Y = y) = \phi I_{(y,0)} + (1 - \phi) g(y; \theta) \quad y = 0, 1, 2, \dots \quad 6.8$$

donde  $I_{(y,0)} = 1$  si  $y = 0$ , y  $I_{(y,0)} = 0$  en otro caso.  $g(y; \theta)$  es la función de probabilidad de la distribución Poisson o de la distribución binomial. Con vector de parámetros  $\theta$ .

Cuando la variable  $X$  sigue una distribución Poisson de parámetro  $\lambda$ , es decir, cuando  $g(y, \lambda) = e^{-\lambda} \lambda^y / y!$ , se tiene un modelo ZIP mientras que cuando la variable  $X$  sigue una distribución binomial con parámetros  $n$  y  $p$ , es decir, cuando  $g(y, n, p) = C_n^y p^y (1 - p)^{n-y}$ , se tiene un modelo ZIB.

### 6.3 Modelo lineal generalizado (MLG)

El modelo lineal generalizado (Nelder y Wedderburn, 1972), puede ser aplicado en casos en que la variable de respuesta no necesariamente tiene distribución normal y para cierto grado de no linealidad en la estructura del modelo.

El MLG se define como;

$$g(\mu_i) = X_i'\beta \quad 6.9$$

donde  $\mu_i = E(Y_i)$ ,  $g$  es una función de suavizado monótono conocida como función link o función de ligazón,  $X_i$  es la  $i$ -ésima fila de la matriz  $X$  y  $\beta$  es el vector de parámetros desconocidos.

En el MLG en lugar de modelar la media,  $\mu_i$ , como función lineal de los predictores,  $x_i$ , se introduce una transformación monótona y diferenciable  $g(\cdot)$  tal que  $\eta_i = g(\mu_i)$ , donde  $g(\cdot)$  es llamada la función de liga y  $\eta_i$  es el predictor lineal asumiendo que la media transformada sigue un modelo lineal,  $\eta_i = X_i'\beta$ . Además, la función link es invertible, es decir;  $\mu_i = g^{-1}(\eta_i) = g^{-1}(X_i'\beta)$ . En el caso canónico se cumple que  $g(\mu_i) = \theta_i = \eta_i = X_i'\beta$ . Cabe mencionar que la función de liga no está transformando los datos en su variable respuesta, sino que transforma el valor esperado,  $\mu_i$ .

El MLG establece además que las  $Y_i$  son independientes y que pertenecen a alguna distribución de la familia exponencial. La familia exponencial de distribuciones incluye muchas distribuciones que son útiles para ser modeladas en la práctica, como por ejemplo, las distribuciones Poisson, binomial, gamma y normal, entre otras. Una referencia importante es Mc Cullagh y Nelder, (1989).

Debido a que el MLG está especificado en términos del “predictor lineal”,  $X_i'\beta$ , muchas de las ideas y conceptos de los modelos lineales (ML) se han extendido a los MLG. La formulación del modelo es la misma, excepto que deben especificarse la función de distribución de la variable de respuesta y la función de ligazón adecuada. En particular, cuando la función de ligazón elegida es la identidad, junto con la distribución normal, el modelo lineal pasa a ser un caso particular de los MLG.

La generalización, sin embargo, tiene un costo adicional en el sentido de que el ajuste del modelo se realiza de manera iterativa y la precisión de los resultados inferenciales depende en gran medida del tamaño de la muestra.

Como ya fue establecido, en el MLG la variable de respuesta puede tener cualquier distribución de la familia exponencial. Una distribución pertenece a la familia de distribuciones exponenciales si su función de densidad de probabilidad puede ser expresada como;

$$f_{\theta}(y) = \exp\{y\theta - b(\theta)\}/a(\varphi) + c(y, \varphi) \quad 6.10$$

donde  $b$ ,  $a$  y  $c$  son funciones arbitrarias,  $\varphi$  es un parámetro de escala y  $\theta$  es conocido como el parámetro canónico de la distribución.

Por otro lado, la esperanza de cualquier variable aleatoria de la familia exponencial está dada por la primera derivada de  $b$ , mediante la relación  $E(Y) = b'(\theta)$ , donde la forma de  $\theta$  depende de la distribución particular. Esta ecuación es la base de la relación del enlace (link) de los parámetros del modelo,  $\beta$ , con los parámetros canónicos de la familia exponencial. En los MLG, los parámetros  $\beta$  determinan la media de la variable de respuesta  $y$ , por lo tanto, también determinan el parámetro canónico para cada respuesta observada.

### 6.3.1 Modelo lineal generalizado aplicado a la distribución binomial

Es sencillo ver que la distribución binomial es un miembro de la familia exponencial. Supóngase que  $Y$  es una variable aleatoria que indica el número de unidades no conformes en una muestra de  $n$  unidades, y que  $y_i/n$  ( $i=1,2,\dots, n$ ) representa la proporción de unidades no conformes en la muestra (es claro que el tamaño de la muestra puede ser variable pero lo consideraremos fijo por simplicidad y debido a que es la situación más frecuente en el control de procesos).

Luego,  $Y$  es una variable aleatoria binomial, tal que  $Y_i \sim \text{Bin}(n, p)$ , donde  $p$  es la probabilidad de éxito, es decir, la probabilidad de encontrar una unidad no conforme en la muestra de  $n$  unidades. Luego,

$$f(y, n, p) = \binom{n}{y} p^y (1-p)^{n-y} = \binom{n}{y} \left(\frac{p}{1-p}\right)^y (1-p)^n \quad 6.11$$

Esta expresión se puede reescribir en su forma canónica para la familia exponencial de la siguiente manera:

$$\begin{aligned} f(y, n, p) &= \exp\left\{y \log\left(\frac{p}{1-p}\right) + n \log(1-p) + \log\left[\binom{n}{y}\right]\right\} \\ &= \exp\left\{y \log\left(\frac{p}{1-p}\right) - n \log\left(\frac{1}{1-p}\right) + \log\left[\binom{n}{y}\right]\right\}. \end{aligned} \quad 6.12$$

donde  $\log\left(\frac{p}{1-p}\right)$  es el parámetro natural de la distribución.

Extendiendo este resultado para cada  $i = 1, 2, \dots, n$ , y considerando una probabilidad de éxito fija y un tamaño de muestra constante, se tiene;

$$f(\mathbf{y}, n, p) = \exp \left\{ \sum_{i=1}^n y_i \log \left( \frac{p}{1-p} \right) - n \log \left( \frac{1}{1-p} \right) + \log \left[ \binom{n}{\sum_{i=1}^n y_i} \right] \right\}. \quad 6.13$$

Como un caso particular de la Ecuación 6.10, donde;

$$\theta(p) = g(\mu) = \log \left( \frac{p}{1-p} \right) \Rightarrow p/(1-p) = e^\theta \Rightarrow p = g^{-1}(\theta) = \frac{e^\theta}{1+e^\theta}$$

Siendo  $g(\mu) = \log \left( \frac{p}{1-p} \right) = \eta$ , conocido como el logit de  $\mu$

Además,

$$b(\theta) = n \log \left( \frac{1}{1-p} \right) = n \log \left( \frac{1}{1-\frac{e^\theta}{1+e^\theta}} \right) = n \log (1 + e^\theta), \quad a(\varphi) = 1 \text{ y } c(\mathbf{y}, \varphi) = \log \left[ \binom{n}{\sum_{i=1}^n y_i} \right].$$

En particular; el valor esperado y la varianza de la distribución binomial se pueden obtener a partir de la familia exponencial como:

$$E(Y) = b'(\theta) = \frac{\partial b(\theta)}{\partial \theta} = \frac{\partial [n \log(1 + e^\theta)]}{\partial \theta} = n \left( \frac{e^\theta}{1 + e^\theta} \right) = np = \mu$$

$$V(Y) = b''(\theta)a(\varphi) = n \left[ \left( \frac{e^\theta}{1 + e^\theta} \right) - \left( \frac{e^\theta}{1 + e^\theta} \right)^2 \right] = np(1 - p) = \frac{\mu}{n}(n - \mu) \quad 6.14$$

Esta última expresión ha sido reescrita como una función explícita de  $\mu$ , para mostrar que, cuando la media de la función binomial es conocida, la varianza queda completamente especificada.

Este hecho es una importante restricción en muchas situaciones prácticas debido a que cuando el modelo binomial se ajusta a datos con sobredispersión, se tiende a subestimar la variabilidad de los datos y, consecuentemente a sobreestimar la significación de los parámetros del predictor lineal. Utilizar el modelo binomial para datos con sobredispersión incrementa las señales de falsa alarma por lo que se hace necesario recurrir a otros modelos para tratar con esta variación extra-binomial (Hinde y Demetrio, 1998).

### 6.3.2 Estimación de los parámetros

Para ajustar un MLG dado por  $g(\mu_i) = X_i' \beta$ , con  $Y_i \sim f_{\theta_i}(y_i)$ , donde  $f_{\theta_i}(y_i)$  indica una distribución de la familia exponencial, con parámetro canónico  $\theta_i$  determinado por  $\mu_i$  vía  $E(Y)$ , es

necesario estimar  $\beta$ . Dado un conjunto de observaciones mutuamente independientes  $Y_i$ , el estimador máximo verosímil de  $\beta$  se define a partir de su verosimilitud como;

$$L(\beta) = \prod_{i=1}^n f_{\theta_i}(y_i) \quad 6.15$$

Y luego el logaritmo de la verosimilitud es;

$$l(\beta) = \sum_{i=1}^n \log [f_{\theta_i}(y_i)] = \sum_{i=1}^n \{y_i \theta_i - b_i(\theta_i)\} / a_i(\varphi) + c_i(y_i, \varphi) \quad 6.16$$

La dependencia del lado derecho de  $l(\beta)$  con  $\beta$  es a través de la dependencia de  $\theta_i$  con  $\beta$ , dada a través de la función liga canónica de la siguiente manera;

$$\theta_i = \eta_i = g(E(Y)) = g(\mu) = X_i' \beta \quad 6.17$$

Después de algunos cálculos y considerando que  $\varphi$  es conocida, se obtiene el estimador de  $\beta$  como aquel valor de que maximiza a la función de verosimilitud, es decir;

$$l(\hat{\beta}) = \max_{\beta} \{l(\beta)\} \quad 6.18$$

En general, en los MLG no existe una expresión cerrada para  $\hat{\beta}$  de tal forma que se requiere de un algoritmo numérico para calcularlo. En este desarrollo se ha empleado el paquete VEGAM del software R que maximiza el logaritmo de la verosimilitud mediante el procedimiento de mínimos cuadrados iterativamente reponderados (IRLS), basado en el algoritmo de Newton-Raphson y la matriz de información de Fisher.

Para encontrar los estimadores de máxima verosimilitud  $\beta$  en el modelo binomial generalizado se define la función de verosimilitud de la siguiente manera;

$$L(n, p) = f(\mathbf{y}, n, p) = \prod_{i=1}^n f(y_i, n, p) = \prod_{i=1}^n \binom{n}{y_i} p^{y_i} (1-p)^{n-y_i} \quad 6.19$$

y tomando el logaritmo,

$$\begin{aligned} l(\mathbf{y}, n, p) &= \sum_{i=1}^n \log \left[ \binom{n}{y_i} p^{y_i} (1-p)^{n-y_i} \right] \\ &= \sum_{i=1}^n \left\{ y_i \log \left( \frac{p}{1-p} \right) - n \log \left( \frac{1}{1-p} \right) + \log \binom{n}{y_i} \right\} \end{aligned} \quad 6.20$$

Pero las ligas canónicas se presentan cuando,

$$\theta = g(\mu) = \text{logit}(\mu) = X'\beta \Rightarrow \hat{p} = g^{-1}(X'\beta) = \frac{e^{X'\hat{\beta}}}{1 + e^{X'\hat{\beta}}} \quad 6.21$$

Luego es posible reescribir el logaritmo como;

$$l(p) = \sum_{i=1}^n \left\{ y_i \log p - y_i \log (1 - p) + n \log(1 - p) + \log \binom{n}{y_i} \right\} \quad 6.22$$

y sustituyendo en la expresión anterior la función liga (6.21), se obtiene;

$$l(\beta) = \sum_{i=1}^n \left\{ \begin{array}{l} y_i [X'_i\beta - \log (1 + e^{X'_i\beta})] - y_i \log \left( 1 - \frac{e^{X'_i\beta}}{1 + e^{X'_i\beta}} \right) + \\ n \log \left( 1 - \frac{e^{X'_i\beta}}{1 + e^{X'_i\beta}} \right) + \log \binom{n}{y_i} \end{array} \right\} \quad 6.23$$

donde  $\beta$  es el vector de parámetros a ser estimados y  $\mathbf{X}$  es un vector de variables explicativas. En este caso se utiliza la función logit como función link.

Luego, el estimador máximo verosímil  $\hat{p}$  de  $p$  se obtiene a partir de la Ecuación 6.21.

En el caso particular de la construcción de un gráfico de control para un proceso binomial, se pueden tener variables explicativas, pero esta no es la situación que se va a considerar en este Capítulo. Aquí se tratará el caso en el que se cuenta con un vector de observaciones de un proceso asociado a una característica de calidad. Por lo tanto, la estimación del parámetro es mucho más simple y se determina de la siguiente manera;

$$\hat{p} = \frac{e^{\hat{\beta}}}{1 + e^{\hat{\beta}}} \quad 6.24$$

Es decir, sólo será considerada la estimación del valor de  $\beta$  para la ordenada al origen (intercept) del modelo lineal.

### 6.3.3 Modelo binomial con exceso de ceros

El modelo con exceso de ceros puede considerarse como un modelo mixto Bernoulli-binomial (Vieira *et al.*, 2000), donde una posible explicación del exceso de ceros en los datos puede provenir de un mejoramiento considerable en alguna característica de calidad bajo monitoreo.

Supóngase que  $Y_i$  es el número de unidades no conformes en la muestra y sea  $Z_i = I_{(0)}$ , donde  $I_{(0)}$  es una función indicadora de la conformidad de la muestra particular  $i$ .

Condicionamente se puede asumir que;

$$\begin{cases} \{Y_i|Z_i = 0\} \sim \text{Binomial}(n, p) \\ \{Y_i|Z_i = 1\} \equiv 0 \end{cases} \quad 6.25$$

donde  $Z_i$  es una variable aleatoria con función de probabilidad Bernoulli, dada por;

$$P(Z = z) = \phi^z (1 - \phi)^{1-z} \quad \text{para } z = 0, 1 \quad 6.26$$

y  $\phi$  es la probabilidad de observar un exceso de ceros.

De esta manera la función de probabilidad de  $Y$  se puede expresar, evitando los subíndices por simplicidad, de la siguiente manera;

$$P(Y = y) = \begin{cases} \phi + (1 - \phi) (1 - p)^n & y = 0 \\ (1 - \phi) \binom{n}{y} p^y (1 - p)^{n-y} & y = 1, 2, \dots, n \end{cases} \quad 6.27$$

que conduce a los momentos,

$$E(Y) = (1 - \phi)np = (1 - \phi)\mu \quad 6.28$$

$$Var(Y) = (1 - \phi)np [(1 - p) + np\phi] = (1 - \phi)\mu \left[ \left( \frac{n-\mu}{n} \right) + \phi \mu \right]$$

El logaritmo de la verosimilitud basado en estos datos está dado por;

$$\begin{aligned} l(\beta, \phi, \mathbf{y}, \mathbf{z}) &= \sum_{i=1}^n \log [f(y_i, z_i; \beta, \phi)] = \sum_{i=1}^n \log [f(y_i|z_i; \beta, \phi)f(z_i|\phi)] \\ &= \sum_{i=1}^n \log \left[ \binom{n_i}{y_i} p^{y_i} (1 - p)^{n-y_i} \right]^{z_i} + \sum_{i=1}^n \log [\phi^{z_i} (1 - \phi)^{1-z_i}] \end{aligned} \quad 6.29$$

Para un modelo binomial simple, con  $n$  y  $p$  constantes para todo  $i$ , y expandiendo la función logaritmo, se tiene;

$$\begin{aligned}
 l(\beta, \phi, \mathbf{y}, \mathbf{z}) &= \sum_{i=1}^n z_i \left[ \log \binom{n}{y_i} + y_i \log \left( \frac{p}{1-p} \right) + n \log(1-p) \right] \\
 &+ \sum_{i=1}^n z_i \log \left( \frac{\phi}{1-\phi} \right) + n \log(1-\phi) = l(\beta; \mathbf{y}, \mathbf{z}) + l(\phi; \mathbf{z})
 \end{aligned}
 \tag{6.30}$$

donde el  $\log \left( \frac{\phi}{1-\phi} \right)$  se conoce como el logit de  $\phi$ .

Esta expresión, muestra que la estimación de  $p$  y de  $\phi$ , se pueden obtener maximizando los dos términos de  $l(\beta, \phi, \mathbf{y}, \mathbf{z})$  separadamente.

En R esta función se conoce como “zibinomial” y los dos predictores lineales/aditivos predeterminados (default) son el logit de  $\mu$  y el logit de  $\phi$ .

#### 6.4 Selección del modelo adecuado

La distribución binomial ha sido ampliamente utilizada para modelar datos de conteo, tales como el número de no conformidades por muestra en procesos productivos. Como ya fue establecido, cuando el proceso opera con alta calidad, es probable que muchas muestras sean conformes es decir que no contengan unidades con defectos y por lo tanto se genera una gran cantidad de ceros en la variable de respuesta.

Ante esta situación, el investigador deberá determinar si el proceso es binomial con un número elevado de ceros debido a una tasa de no conformidades muy chica, o si se está en presencia de una cantidad de ceros que es superior a la esperada para una distribución binomial estándar. En pocas palabras, se debe decidir entre una distribución binomial estándar y una distribución binomial con exceso de ceros.

Esta decisión es de gran importancia, teniendo en cuenta que la habilidad de un gráfico de control está determinada por la bondad de ajuste entre los datos observados y la distribución asumida para la construcción del gráfico. Cuando se fracasa en tener en cuenta el exceso de ceros, se genera un modelo mal especificado que puede resultar en estimadores sesgados o inconsistentes.

La bondad de ajuste entre modelos competitivos puede establecerse utilizando varios criterios de información estadística. En particular, en los MLG estas estadísticas son de vital importancia para determinar la distribución de probabilidad que mejor ajusta a los datos. Entre los criterios más populares se encuentran el Criterio de Información de Akaike (AIC) (1974), una extensión del AIC conocida como Criterio AICC (Hurvich y Tsai, 1989) y el Criterio de Información Bayesiana (BIC)

(Schwarz, 1978). Todos estos criterios se basan en la función de verosimilitud corregida por la cantidad de parámetros a estimar.

En este trabajo, se empleará el Criterio de Akaike que permite evaluar tanto el ajuste del modelo a los datos como la complejidad del modelo. Cuanto más pequeño es el valor del AIC mejor es el ajuste. El AIC es útil para comparar modelos similares con distintos grados de complejidad.

El Criterio de Akaike se define como:

$$AIC = 2[-l(\hat{\beta}) + d] \quad 6.31$$

donde  $l(\hat{\beta})$  es el valor máximo del logaritmo de la verosimilitud y  $d$  es el número de parámetros en el modelo cuando las observaciones son independientes.

## 6.5 Determinación de los límites de control

Una vez que se ha determinado el modelo adecuado a los datos, se pueden construir los límites de control. Si la distribución elegida es la binomial, es posible emplear cualquiera de los gráficos desarrollados en los Capítulos anteriores. Ahora será considerado el gráfico adecuado para el caso en el que los datos presentan un exceso de ceros es decir, cuando el modelo seleccionado es un modelo ZIB.

Generalmente el límite inferior de control para gráficos basados en el modelo ZIB no se considera. Dado que la distribución ZIB es altamente asimétrica, para definir los límites de control se sugiere emplear límites de probabilidad. Luego, con una señal de falsa alarma predeterminada  $\alpha$ , el límite superior  $k$  de control del gráfico basado en el número de no conformidades se obtiene como el menor valor entero que satisface;

$$P(\text{de } k \text{ o más no conformidades en la muestra}) \leq \alpha \quad 6.32$$

Sin embargo, se debe notar que, como la distribución ZIB es discreta, no es siempre posible especificar el nivel de confianza con exactitud, dado que  $k$  sólo puede tomar valores enteros. En esta situación conviene tomar el límite superior de control como aquel valor de  $k$  para el cual:

$$LSC = \min \left\{ k: \phi + (1 - \phi)(1 - p)^n + \sum_{y=1}^{k-1} (1 - \phi) \binom{n}{y} p^y (1 - p)^{n-y} > 0,9973 \right\} \quad 6.33$$

## 6.6 Análisis de sensibilidad del gráfico de control

Es importante estudiar la sensibilidad del gráfico de control basado en el modelo ZIB. Para ello se empleará la longitud promedio de corrida, ARL. Como el modelo ZIB contiene dos parámetros, cada uno puede tener un impacto diferente sobre la probabilidad de alarma. Aquí será considerado sólo el límite superior del gráfico y se analizará el efecto de los cambios de cada uno de los parámetros del modelo suponiendo un proceso estable.

Como ya fue determinado, el ARL se obtiene mediante la relación  $ARL = 1/(1 - \beta)$ , es decir, la inversa de la potencia estadística de la prueba siendo  $\beta$  una medida de la habilidad del gráfico de control para detectar desplazamientos en los parámetros del proceso. Cuando  $p$  toma el valor  $p_0$ , que indica el nivel aceptable de calidad o el nivel medio de calidad estimado del proceso, el  $ARL = 1/\alpha$ , donde  $\alpha$  es la probabilidad de falsa alarma.

Para un modelo ZIB, la fórmula general para  $\beta$ , es;

$$\beta = P(D < LSC | \phi, p, \alpha) = \phi + \sum_{y=0}^{LSC-1} (1 - \phi) \binom{n}{y} p^y (1 - p)^{n-y} \quad 6.34$$

donde  $D$ , el número de no conformidades en la muestra, es una variable ZIB con parámetros  $\phi$  y  $p$ , y  $\alpha$  es la probabilidad del error de Tipo I.

Como ya se ha determinado en Capítulos anteriores, la longitud promedio de corrida, ARL, indica el promedio de muestras necesarias hasta obtener un punto fuera del límite de control. Para el modelo ZIB, este valor está dado por;

$$ARL = \frac{1}{P(\text{un punto fuera de control})} = \frac{1}{\sum_{y=LSC}^{\infty} (1 - \phi) \binom{n}{y} p^y (1 - p)^{n-y}} \quad 6.35$$

En la Tabla N° 6.1 se muestran algunos valores del ARL para un proceso con  $p_0 = 0,010$  y  $\phi_0 = 0,5$ , considerando diferentes desplazamientos en los valores de  $p$  y de  $\phi$  y muestras de tamaño 100, 200, 300 y 400.

Tabla 6.1: Valores de ARL para  $p_0 = 0,0100$ ,  $\phi_0 = 0,5$  y  $\alpha = 0,0027$

Límite superior de control	$n$	$p$	$\phi = 0,1$	$\phi = 0,3$	$\phi = 0,5$	$\phi = 0,7$	$\phi = 0,9$
	100	0.005	7005	9007	12610	21017	63052
LSC = 5		<b>0.010</b>	323	416	<b>582</b>	971	2913
		0.015	62	80	113	188	565
		0.020	21	28	39	65	196
	200	0.005	14428	18550	25970	43284	129854
LSC = 7		<b>0.010</b>	258	332	<b>465</b>	775	2327
		0.015	34	44	61	102	308
		0.020	10	13	18	30	92
	300	0.005	43577	56028	78440	130733	392200
LSC = 9		<b>0.010</b>	308	396	<b>555</b>	925	2775
		0.015	28	36	51	85	256
		0.020	7	9	13	22	66
	400	0.005	25697	33039	46254	77091	231273
LSC = 10		<b>0.010</b>	142	183	<b>256</b>	427	1281
		0.015	13	17	24	40	121
		0.020	3	5	7	11	35

Debe notarse que como el límite superior de control es estrictamente entero, no es siempre posible especificar el valor del ARL con exactitud. Esto produce algunas diferencias en los valores observados en la Tabla 6.1 respecto al los valores reales del ARL.

El análisis de sensibilidad indica que este gráfico es relativamente sensible a aumentos en el valor de  $\phi$ . Esto es importante en el control estadístico de procesos dado que  $\phi$  está relacionado con la presencia de unidades conformes. Cuando  $\phi$  aumenta, hay más unidades sin defectos, y luego la cantidad de unidades que deben ser inspeccionadas antes que se dé una señal de alarma aumenta rápidamente para cualquier tamaño de muestra. Por otro lado, a medida que la fracción de no conformes  $p$  aumenta, alejándose del valor de la hipótesis nula, aumenta la potencia y disminuye la cantidad de muestras a ser inspeccionadas antes de dar la alarma.

## 6.7 Aplicación del modelo ZIB a un proceso industrial

Una empresa autopartista fabrica un tornillo de acero en forma de U, conocido técnicamente como tornillo U-bolt. Este tornillo tiene rosca en los extremos y se utiliza para la sujeción de elásticos en ciertos vehículos. Por ser un componente de seguridad de los automóviles el nivel de defecto aceptado debe ser muy bajo. Diariamente la empresa produce 1100 tornillos en cada uno de tres turnos de trabajo por lo que la producción diaria es de 3300 tornillos.

La característica crítica a controlar está relacionada con las posibles fisuras en la estructura del tornillo. Esta característica es de vital importancia dado que el tornillo sujeta una parte del vehículo

de gran importancia para la seguridad de conducción. El control se realiza aplicando una tinta penetrante Magnaflux a cada tornillo que luego es observado mediante una luz fluorescente a efectos de establecer la presencia de fisuras en el material. Dada la criticidad de la característica, debería realizarse un control del 100%, pero no es posible hacerlo debido a que el proceso de control insume demasiado tiempo. Se ha implementado un sistema de control del proceso que consiste en seleccionar 200 tornillos cada tres horas de producción y someterlos al control de fisuras.

Debido a que se trata de un proceso nuevo, el Dpto. de Calidad se encuentra realizando ajuste al mismo que permitan reducir la cantidad de piezas con defectos de fisura. Para estudiar el comportamiento y la evolución del proceso se ha decidido implementar un gráfico de control para la cantidad de piezas con fisuras verificadas en cada muestra.

Cabe destacar que el proceso de control no es destructivo. Las partes sin fisuras se devuelven a la producción y las que tienen fisuras son investigadas con mayor atención a efectos de determinar la posible causa de la falla y tomar las acciones correspondientes.

En la Figura N° 6.1 se muestra el tipo de vehículo para el cual se fabrica el tornillo U-bolt, el tornillo y la localización de la pieza bajo estudio.



Figura 6.1: Tipo de vehículo y localización de la pieza bajo estudio

Este caso será utilizado para ilustrar el modelo ZIB. La bondad de ajuste de este modelo se compara con el modelo binomial a efectos de establecer el más adecuado para los datos del problema. El estudio se complementará además con otros procedimientos estadísticos que pueden ser de utilidad a la hora de tener que especificar el modelo adecuado para un conjunto de datos con exceso de ceros.

Los datos que se presentan a continuación corresponden a los controles realizados sobre 260 muestras de 200 tornillos cada una. En primer lugar, se presenta la Tabla 6.2 en la que se muestra la distribución de la cantidad observada de tornillos con fisuras.

Tabla 6.2: Tabla de frecuencia

Cantidad de tornillos con fisuras por muestra	Cantidad de muestras
0	154
1	63
2	29
3	8
4	4
5	1
6	1

Las frecuencias observadas parecen indicar que la cantidad de muestras en las que no se encontraron tornillos con fisuras tiene una frecuencia más elevada que la que se debería esperar si la muestra hubiera sido generada mediante una distribución binomial. Para comprobarlo, el paso siguiente consistió en emplear el modelo lineal generalizado a efectos de determinar si el proceso es binomial con un número elevado de ceros o si se está en presencia de una cantidad de ceros que es superior a la esperada para una distribución binomial estándar, generando un modelo ZIB. La Tabla 6.3 muestra el modelo que mejor ajusta los datos, recordando que, cuanto más pequeño es el valor del AIC, mejor es el ajuste. También se encuentran los parámetros estimados para ambos modelos.

Tabla 6.3: Elección del modelo y estimación de los parámetros

Modelo	Criterio AIC de Akaike	$\phi$	$\hat{p}$
Binomial	605,6207	0	0,00331
ZIB	587,1644	0,380	0,00532

Los resultados indican que el modelo ZIB es el modelo adecuado para estos datos.

A efectos de visualizar la importancia de la correcta especificación del modelo, se calcularon las frecuencias relativas observadas y las probabilidades estimadas para cada uno de los modelos propuestos. Debido a que las frecuencias esperadas para valores de la variable mayores o iguales a 3 son muy pequeñas se las ha agrupado en una categoría que se ha llamado 3+. Estos valores se encuentran en la Tabla 6.4 y están graficados en la Figura 6.2.

Tabla 6.4: Probabilidades estimadas para los modelos binomial y ZIB

Valores	Frecuencia absoluta	Frecuencia relativa	Probabilidad binomial	Probabilidad ZIB
0	154	0,59231	0,42231	0,59334
1	63	0,24231	0,36464	0,22821
2	29	0,11154	0,15682	0,12145
3+	14	0,05385	0,05620	0,05695

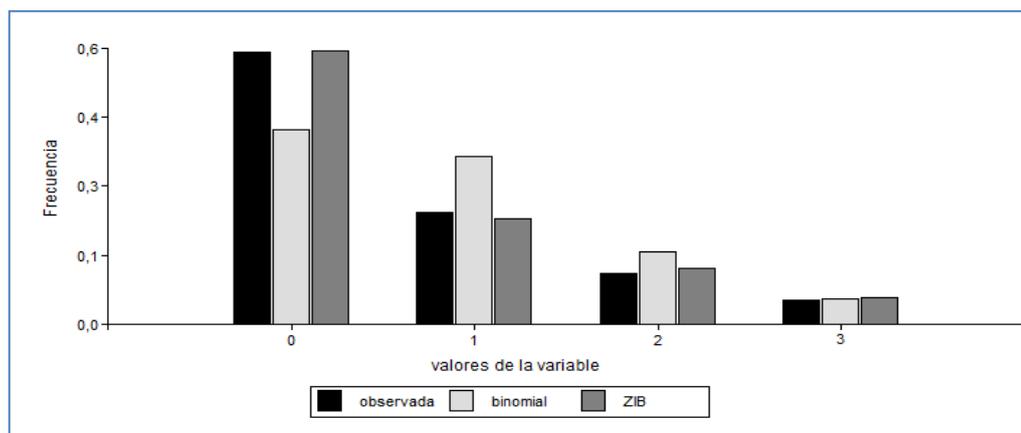


Figura 6.2: Probabilidades observadas y estimadas para los modelos binomial y ZIB

Nótese que el modelo ZIB es mejor para tratar con la probabilidad de  $Y = 0$  y generalmente sigue a la distribución observada mucho mejor que el modelo binomial.

Para establecer el ajuste del modelo también es posible emplear el test de bondad de ajuste. Los resultados del test chi-cuadrado de Pearson indicaron que la distribución binomial no ajusta los datos ( $\chi^2 = 31,9$ ;  $p < 0,0001$ ), mientras que el modelo ZIB los ajusta correctamente ( $\chi^2 = 0,48$ ; ns).

Otro aspecto importante tratado en este Capítulo es el referido al hecho de que cuando el modelo binomial se ajusta a datos con sobredispersión se tiende a subestimar la variabilidad de los datos (Sección 6.3.1). El ejemplo que se está considerando es un caso evidente de datos con sobredispersión por exceso de ceros. El cálculo del promedio y varianza para cada uno de los modelos considerados se muestra en la Tabla 6.4.

Tabla 6.5: Valor esperado y variabilidad para los modelos binomial y ZIB

Modelo	$E(Y)$	$Var(Y)$
Binomial	0,662	0,660
ZIB	0,660	0,923

Una vez que se ha seleccionado el modelo adecuado se puede establecer el límite de control del gráfico. Para los datos de este problema, el límite superior de control es 5 (para el modelo binomial es 4) considerando una tasa de falsa alarma de 0,0027. Esto significa cualquier muestra que tenga 5 o más unidades con defecto en la muestra de 200 unidades, debe ser rechazada.

El gráfico de control para el problema de fisura en los tornillos obtenido sobre 260 muestras de tamaño 200, con un modelo ZIB con parámetros estimados  $\hat{\phi} = 0,38$  y  $\hat{p} = 0,00532$  es el de la Figura 6.3

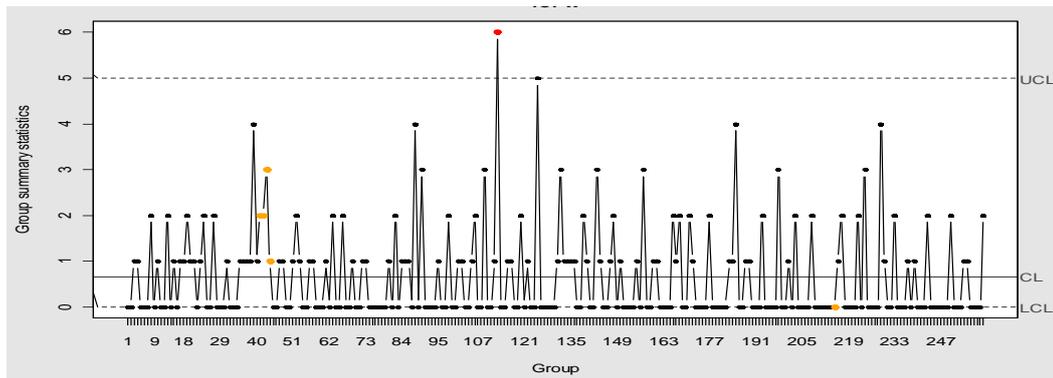


Figura 6.3: Gráfico de control para la cantidad de tornillos con fisuras.

Aunque el gráfico de control indica que sólo hay dos muestras fuera de control, la cantidad de fisuras del proceso debería ser mucho menor. El Dpto. de Calidad deberá tomar acciones correctivas.

## 6.8 Comentario final

Cuando se analizan datos de procesos en los que se observan muchas muestras conformes, es decir, muchas muestras con cero defectos, la utilización del modelo ZIB puede ser una muy buena alternativa. Esto es especialmente cierto para el caso en que los datos muestran una frecuencia más elevada de unidades conformes (cero defectos) que la que se debería esperar si la muestra hubiera

sido generada mediante una distribución binomial. Tradicionalmente estos procesos fueron monitoreados empleando la distribución binomial pero, bajo estas circunstancias, la distribución binomial tiende a subestimar la variabilidad del proceso. En esta situación, los gráficos de control tienen límites muy estrictos que determinan excesivas señales de falsa alarma, altos costos de inspección y frecuentes paradas del proceso. La utilización del modelo lineal generalizado ha mostrado ser un procedimiento adecuado para establecer la bondad de ajuste entre los datos observados y la distribución asumida para la construcción del gráfico. Cuando no se tiene en cuenta el exceso de ceros, se genera un modelo mal especificado y, en consecuencia, el gráfico de control resultante no cumple con la función para la cual ha sido construido.

## Capítulo 7

### Conclusiones, aportes de la investigación y trabajos futuros

#### 7.1 Resultados y Conclusiones

El mejoramiento continuo es considerado hoy en día como la esencia de la supervivencia y crecimiento de una empresa que requiere de una permanente reducción en el nivel de no conformidades de sus procesos. Sin embargo, se ha prestado mucha atención a la aplicación de programas de gestión que permitan a las empresas mejorar sus procesos, pero se ha puesto poca atención en los procedimientos estadísticos necesarios para que un proceso o producto se mantenga sin defectos. En este aspecto, el Control Estadístico de Procesos (CEP) sigue desempeñando un rol fundamental en el esfuerzo hacia el mejoramiento continuo. Controlar y/o asegurar la conformidad del producto con las especificaciones en niveles cercanos a cero defectos, requiere de una reevaluación de ciertos procedimientos estadísticos tales como los gráficos de control.

Los gráficos de control constituyen una herramienta de monitoreo de procesos ampliamente utilizada especialmente en la industria. Ellos permiten detectar desviaciones significativas de las características del proceso debidas a causas asignables. Es común recomendar que cuando se observa una señal de fuera de control, el proceso sea detenido y se investigue esta causa. Este punto de vista tradicional se aleja de la concepción moderna que enfatiza el mejoramiento continuo del proceso como una tarea permanente mediante la cual el proceso evoluciona hacia la mejora. Es necesario trabajar sobre el proceso buscando la reducción permanente del nivel de unidades no conformes más que buscando la causa de una no conformidad.

Teniendo en cuenta estas consideraciones esta tesis se basó en la necesidad de investigar nuevos procedimientos para el monitoreo y control estadístico de atributos en procesos de alta calidad, en función de la magnitud de los parámetros, el diseño de la muestra y las medidas de desempeño.

A continuación se presentan las principales conclusiones del estudio realizado en este trabajo de investigación.

##### **7.1.1 Desarrollo de un gráfico $p$ mejorado para procesos con bajo número de no conformidades y muestras chicas.**

Cuando los procesos tienden a mejorar, el valor esperado de no conformidades se torna cada vez más pequeño y los límites de control basados en la aproximación normal ya no pueden describir

adecuadamente la situación real del estado del proceso. De esta manera se incrementa la probabilidad de falsas alarmas y las señales de que el proceso se encuentra fuera de control pueden ser incorrectamente interpretadas.

Para tratar con esta problemática, surgieron en la literatura varias alternativas que fueron presentadas y analizadas en el Capítulo 3 de la tesis. Una de ellas propone la utilización de límites probabilísticos que tienen en cuenta la distribución subyacente de la característica a controlar. Sin embargo, aunque este procedimiento es estadísticamente correcto, los límites de probabilidad no constituyen una herramienta habitual en el control de procesos debido a que requieren de cierto conocimiento acerca de las distribuciones de probabilidad. Además por lo general, estos límites no son simétricos, lo que constituye un problema práctico dado que dificulta la interpretación del gráfico.

Otra alternativa se orientó hacia la modificación de los datos a efectos de lograr transformar una distribución asimétrica en una casi simétrica. Se presentaron varios enfoques estadísticos que han sido aplicados tanto a la distribución binomial como a la distribución de Poisson, de manera de obtener la transformación que mejor aproxime a la distribución normal. Sin embargo, como se ha establecido en este Capítulo, los procedimientos de transformación de datos para normalización no fueron muy bien recibidos por los responsables de calidad. Ello se debió principalmente a la dificultad relacionada con la transformación en sí misma (cambio de escala de los datos) y con la interpretación de las señales de inestabilidad y de seguimiento de los procesos.

Ante la necesidad de buscar una solución, surgió una propuesta consistente en establecer una modificación, no ya en los datos, sino en los límites de control. Se determinó que una buena opción podía consistir en la utilización de la expansión de Cornish-Fisher. La bibliografía consultada establecía que hasta el presente se había realizado el cálculo de la expansión simple de Cornish-Fisher hasta el cumulante de 3° orden, obteniendo límites de control con un ajuste. Esta modificación permitió corregir el sesgo y la asimetría de la distribución binomial para casos en que la fracción de no conformes  $p$  es chica. Con la incorporación de este simple ajuste es posible aplicar el gráfico  $p$  tradicional, basado en la distribución normal, efectuando sólo una modificación sencilla en los límites de control.

Sobre la base de este resultado, se decidió desarrollar un nuevo gráfico  $p$ , llamado gráfico  $p$  mejorado, para ser utilizado en el control de procesos con valores de  $p$  más pequeños. Para la obtención de los límites modificados se partió de la expansión cuantílica de Cornish-Fisher utilizando la información de los cumulantes de 3° y 4° orden. Esta modificación permitió corregir no sólo el sesgo y la asimetría de la distribución binomial, sino también de la curtosis. La incorporación de un nuevo término de ajuste en los límites de control mostró algunas ventajas

principalmente en el sentido de posibilitar el monitoreo de procesos con una fracción de no conformes más pequeña, como es el caso de procesos de muy alta calidad. Al mismo tiempo se mostró la ganancia significativa en términos de reducir el exceso de falsas alarmas respecto al gráfico  $p$  tradicional. Se trabajó sobre la obtención de la expresión matemática para el cálculo del riesgo alfa y su implementación en R, considerando los diferentes límites obtenidos: metodología tradicional de Shewhart (normal), límites con un término de corrección y límites con dos términos de corrección en la expansión de Cornish Fisher. Los resultados basados en estos estudios fueron presentados en tablas y gráficos en función del tamaño de la muestra y demás parámetros, comparándolos con los de la literatura. La aplicación a un problema industrial permitió ilustrar el desempeño comparativo de los tres gráficos  $p$  desarrollados en este Capítulo y mostrar que el gráfico  $p$  mejorado (con dos términos de corrección) tiene un riesgo de falsa alarma muy cercano al prefijado. Los resultados logrados permitieron determinar que esta es la mejor opción para procesos binomiales cuando la fracción de no conformes es pequeña y el tamaño de muestra también, situación que se ha tornado habitual actualmente en muchas empresas.

Este nuevo procedimiento, basado en la determinación de límites de control con dos términos de corrección, constituye uno de los aportes originales de la tesis. La ventaja práctica asociada con esta propuesta radica en el hecho de que el usuario puede insertar estos nuevos límites de control en un software estadístico de una manera muy sencilla.

Cabe destacar que la implementación del gráfico  $p$  mejorado y la evaluación de los resultados en términos de calidad y costos, dieron lugar a la realización de una tesis de Maestría en Tecnología de los Alimentos. El procedimiento de control propuesto posibilitó dar solución a un problema de calidad de una importante empresa del medio.

### **7.1.2 Construcción de una regla empírica para determinar el gráfico $p$ adecuado.**

En el Capítulo 3 se mostró que, cuando el proceso opera con valores de  $p$  muy bajos (y  $n$  también es chico), una buena práctica para mantener el riesgo de falsa alarma en su valor preestablecido, consiste en trabajar con límites de control modificados. En general existe acuerdo en que cuando  $np(1-p) \geq 5$  y  $p > 0,10$ , es conveniente utilizar la aproximación normal para calcular los límites de control del gráfico  $p$ . A efectos de extender estos resultados al caso en que  $p$  es chico, se efectuó un estudio considerando la modificación en los límites de control del gráfico  $p$  mediante la incorporación de uno y de dos términos de corrección. Con estos resultados y tomando diferentes valores de  $n$ , se construyó la Tabla 3.2. En ella se muestra el mínimo valor de  $p$  que puede ser utilizado a fin de mantener el riesgo de falsa alarma bajo control, con: a) gráficos basados en la distribución normal (sin corrección), b) Cornish-Fisher con un término de corrección y c) Cornish-Fisher con dos términos de corrección. Los resultados obtenidos permitieron lograr una

importante regla empírica mediante la cual es posible ahora efectuar la elección apropiada de los límites de control para el gráfico  $p$ , teniendo en cuenta el tamaño de la muestra y el valor de la fracción de no conformidades del proceso.

Cabe destacar que todos los cálculos efectuados con esta investigación fueron realizados mediante el desarrollo de un programa propio elaborado con el software R, el que se encuentra en el Anexo III.

Los resultados logrados con el trabajo de investigación, fueron publicados en el Journal of Control Engineering Practice bajo el título: “An improved attribute control chart for monitoring non-conforming proportion in high quality processes”.

### **7.1.3 Extensión de los gráficos de control con muestreo doble al caso de procesos con baja proporción de no conformes y muestras chicas.**

Como complemento de la investigación abordada en la tesis se realizó el estudio de los gráficos de control con muestreo doble (MD). El MD es un procedimiento que permite mejorar la capacidad de observar cualquier condición de fuera de control mediante la obtención de una segunda muestra. Además permite detectar cambios pequeños o moderados en la proporción de unidades no conformes del proceso, sin aumentar el muestreo.

En el año 2011 Rodrigues *et al.*, presentaron un trabajo en el que propusieron un muestreo en dos etapas empleando los gráficos de control de Shewhart en la primera etapa del muestreo. Sin embargo este procedimiento es válido sólo para muestras grandes. Pensando que en muchas situaciones prácticas, especialmente en la industria, es más común emplear muestras pequeñas que grandes, se determinó la necesidad de desarrollar una metodología que tengan en cuenta esta situación.

Por este motivo, en el Capítulo 4 los resultados obtenidos por Rodrigues *et al.*, se extendieron al caso de procesos con baja proporción de no conformidades y tamaños de muestra chicos. Esta nueva versión del MD fue desarrollada empleando el gráfico  $np$  mejorado (presentado en el Capítulo 3) en lugar del gráfico tradicional de Shewhart, para la primera etapa del muestreo. Se estableció el procedimiento de construcción del gráfico con MD y el plan de control, determinando las diferentes acciones posibles. Posteriormente se definieron las medidas de desempeño del gráfico y el cálculo del tamaño promedio de muestra que depende de si una segunda muestra es necesaria o no. Se mostró la ventaja de emplear en la primera etapa del muestreo (MS) el gráfico  $np$  mejorado respecto al gráfico  $np$  basado en la distribución normal, para el caso  $n$  y  $p$  chicos, mediante el cálculo del porcentaje de ganancia del ARL bajo control.

El paso siguiente consistió en diseñar y aplicar este nuevo procedimiento de MD. Se proporcionaron diseños que minimizan el ARL para incrementos del 50%, 100% y 200% en la proporción de no conformidades del proceso. Este procedimiento fue aplicado a procesos con una fracción de unidades no conformes bajo control ( $p_0$ ) de 0,005, 0,010 y 0,020 y para  $ARL_0 \geq 200$  y  $ARL_0 \geq 370,4$ . Además se mostraron los diseños óptimos para tamaños muestrales promedio de manera que  $np_0 = 0,5, 1,0, 2,0$  y  $4,0$ .

Para resolver el problema de optimización se desarrolló un programa en R mediante el cual se efectuaron los cálculos que se presentaron en las Tablas 4.2 a 4.5. En ellas es posible observar los proyectos óptimos para los gráficos  $np$  con MS y MD y las medidas de rendimiento para todas las combinaciones de valores dados.

A partir de los resultados obtenidos se pudo establecer que el procedimiento de control con MD es especialmente útil cuando se lo utiliza en procesos de alta calidad. En esta situación, la alternativa de emplear el gráfico  $np$  mejorado en la primera etapa del muestreo, ha mostrando una gran ventaja sobre el gráfico  $np$  habitual para atributos. Complementando la primera etapa del muestreo con la utilización del MD en una segunda etapa, permite obtener dos ventajas adicionales; reducir la cantidad total de inspección y darle al proceso una segunda oportunidad antes de tomar una decisión.

Los resultados de esta investigación fueron publicados en el Journal of Statistical Methodology bajo el título “Extending a double sampling control chart for non-conforming proportion in high quality processes to the case of small samples”.

#### **7.1.4 Estudio de los gráficos de control para procesos cercanos a cero defectos, revisión de la metodología y propuestas de aplicación.**

En los últimos años se ha puesto gran atención en el área de control de procesos en los gráficos de control basados en el conteo acumulado de unidades conformes. Estos gráficos se realizan sobre la base de contar la cantidad de unidades conformes producidas entre unidades no conformes.

En el Capítulo 5, se efectuó la revisión de un conjunto de propuestas referidas a la implementación de este tipo de gráficos de control, entre los que se consideraron los gráficos CCC, CCC- $r$ , CCS y CQC. Los gráficos CCC constituyen un procedimiento de control adecuado para procesos con muy baja tasa de no conformidades, en los que las unidades producidas son inspeccionadas una por una, es decir secuencialmente, siguiendo el orden de producción. La variable de interés es el número acumulado de unidades conformes entre no conformes. Sin embargo, cuando la tasa de producción es muy elevada, el conteo es excesivo y dificulta el

procedimiento de control. En esta situación conviene emplear el gráfico CCC- $r$  que realiza el conteo hasta detectar  $r$  no conformidades.

No obstante, aunque estos gráficos han demostrado ser útiles en el seguimiento de procesos de alta calidad, presentan un ARL sesgado, es decir, el valor máximo del ARL, como función de  $p$ , no se encuentra en el nivel nominal del proceso. Esto implica en la práctica que la señal de alarma puede tomar más tiempo en manifestarse cuando el proceso está deteriorado que cuando está bajo control y mientras tanto el proceso será considerado bajo control.

Para evitar esta dificultad, Xie *et al.* (2000, 2002), propusieron un ajuste en los límites de control del gráfico CCC mediante un procedimiento de optimización que permite obtener el máximo ARL en el valor nominal de  $p$ . Para el caso de los gráficos CCC- $r$  se encontraron en la literatura dos propuestas para obtener un ARL insesgado. En este Capítulo se analizaron los fundamentos teóricos de ambos procedimientos. Además, mediante un estudio de simulación, se compararon las curvas de ARL para ambas propuestas del gráfico CCC- $r$  con  $r = 2, 3$  y  $4$ , considerando los límites de control para un proceso con  $p_0$  y  $\alpha$  dados. El estudio comparativo reveló que aunque ambos procedimientos tienen un ARL insesgado, uno es más apropiado para la detección de un aumento en la fracción de no conformes del proceso.

En este Capítulo también fueron considerados los gráficos CCS. Estos gráficos se desarrollaron para procesos en los que las unidades producidas son inspeccionadas por muestras siendo la probabilidad de que una muestra contenga más de un producto no conforme muy pequeña. El problema de estos gráficos está relacionado con el hecho de que no se puede establecer que sucede con el proceso en las muestras que no fueron observadas. Esto dificulta la interpretación correcta de las señales de inestabilidad del proceso. Zhang *et al.*, publicaron varios trabajos referidos a la construcción de gráficos para monitorear el número acumulado de muestras hasta obtener un número especificado de unidades no conformes, que naturalmente incluye a los gráficos CCC y CCC- $r$  como casos especiales. Al igual que estos últimos, el gráfico CCS también mostró problemas de sesgo del ARL por lo que en un trabajo reciente Zhang *et al.* (2012) propusieron una versión mejorada sólo para el caso  $r = 1$ . Además establecieron un procedimiento para determinar los límites de control a partir de la maximización del ARL cuando el proceso se encuentra bajo control. Sin embargo, como fue mostrado en el Capítulo 5, el procedimiento presentado se transforma en una metodología sumamente compleja, aún empleando programas de computación apropiados, tanto para su aplicación práctica como para la interpretación de las señales de inestabilidad.

Para analizar e implementar los gráficos presentados en este Capítulo, se desarrolló un programa en R bajo el nombre de gráficos qcc- $r$ . Este programa permite obtener: a) el gráfico CCC de Xie *et*

al. (2001), b) el gráfico CCC- $r$  de Chen y Cheng (2010), c) el gráfico CCC- $r$  de J. T. Chen (2009) y d) el gráfico CCS de Zhang (2005). Este programa se encuentra en el Anexo IV. Consiste en la modificación de la función qcc (del paquete qcc) para que admita gráficos CCC- $r$  mediante la función llamada ahora qcc- $r$ . La importancia de este programa radica en que cualquier usuario interesado en aplicar estos gráficos pueda acceder a ellos de manera muy sencilla.

Para completar el estudio de los gráficos tipo CCC, se presentó el gráfico CQC, basado en la distribución exponencial, que se utiliza en situaciones en que la cantidad producida se mide en una escala numérica más que como un conteo.

Los resultados de esta investigación fueron enviados al Journal of Process Control bajo el título “Study of the CCC- $r$  Control Chart for High-Quality Processes and its Application to Real Data from a Car Spare Parts Plant”.

### **7.1.5 Desarrollo de una nueva metodología para implementar gráficos de control en procesos binomiales con exceso de ceros.**

En el Capítulo 6 se desarrolló una nueva metodología para implementar gráficos de control en procesos binomiales con exceso de ceros. Esta es una característica de los procesos de alta calidad en los que es común observar muchas muestras conformes que generan una gran cantidad de ceros en la variable de respuesta. Ante esta situación, es muy importante determinar si el proceso es binomial con un número elevado de ceros debido a una tasa de no conformidades muy chica, o si se está en presencia de una cantidad de ceros que es superior a la esperada para una distribución binomial estándar. Esta decisión es de gran importancia, teniendo en cuenta que la habilidad de un gráfico de control está determinada por la bondad de ajuste entre los datos observados y la distribución asumida para la construcción del gráfico. Cuando no se tiene en cuenta el exceso de ceros, se genera un modelo mal especificado que puede resultar en estimadores sesgados o inconsistentes.

Para la elección adecuada del modelo se empleó el modelo lineal generalizado por ser éste un procedimiento muy flexible para tratar con datos de conteos. En primer lugar, se presentaron las características de los modelos ZIP (Poisson con exceso de ceros) y ZIB (binomial con exceso de ceros). Posteriormente se definieron las características metodológicas del modelo lineal generalizado y su aplicación al caso binomial y binomial con exceso de ceros, por ser este el procedimiento menos tratado en la literatura, especialmente para el caso del control estadístico de la calidad. En particular, se mostró que el modelo binomial con exceso de ceros se puede modelar como un modelo mixto Bernoulli-binomial que permite obtener la estimación de los parámetros  $p$  y  $\phi$  de manera independiente. Para determinar la bondad de ajuste entre modelos competitivos se

empleó el criterio de Akaike que permite evaluar tanto el ajuste del modelo a los datos como su complejidad.

Basado en estas consideraciones teóricas, se desarrolló el gráfico de control para el caso en que el modelo seleccionado es un modelo ZIB. La sensibilidad del gráfico de control fue evaluada empleando la longitud promedio de corrida, ARL. Como el modelo ZIB contiene dos parámetros, cada uno puede tener un impacto diferente sobre la probabilidad de alarma. Por este motivo, se analizó la sensibilidad del gráfico para diferentes valores de los parámetros del modelo suponiendo un proceso estable. Los resultados obtenidos mostraron que este gráfico es relativamente sensible a aumentos en el valor de  $\phi$ . Esto es importante en el control estadístico de procesos dado que  $\phi$  está relacionado con la presencia de unidades conformes. Cuando  $\phi$  aumenta, hay más unidades sin defectos, y luego la cantidad de unidades que deben ser inspeccionadas antes que se dé una señal de alarma aumenta rápidamente para cualquier tamaño de muestra. Por otro lado, a medida que la fracción de no conformes aumenta, alejándose del valor de la hipótesis nula, aumenta la potencia y disminuye la cantidad de muestras a ser inspeccionadas antes de dar la alarma.

El estudio efectuado en este Capítulo ha permitido establecer que la utilización del modelo lineal generalizado es un procedimiento adecuado para establecer la bondad de ajuste entre los datos observados y la distribución asumida para la construcción del gráfico. Cuando no se tiene en cuenta el exceso de ceros, se genera un modelo mal especificado y, en consecuencia, el gráfico de control resultante no cumple con la función para la cual ha sido construido.

## 7.2 Aportes originales de esta investigación

En esta sección se hace un resumen de los aportes originales que fueron realizados durante esta investigación. Ellos son:

- i) Desarrollo e implementación de un nuevo gráfico de control para procesos con baja proporción de no conformidades y muestras chicas.
- ii) Construcción de una regla empírica a efectos de determinar los límites de control adecuados para procesos que han mejorado mostrando una reducción en la fracción de no conformes:
  - ✓ cuando  $np(1 - p) \geq 5$  sin corrección (normal)
  - ✓ cuando  $np(1 - p) \geq 0,25$  un término de corrección
  - ✓ cuando  $np(1 - p) \geq 0,08$  dos términos de corrección
- iii) Desarrollo de un gráfico de control con muestreo doble para procesos con baja proporción de no conformidades y muestras chicas.

- iv) Determinación del gráfico de control más adecuado para procesos con muy baja proporción de no conformidades, especialmente procesos automatizados, donde las unidades producidas son inspeccionadas secuencialmente siguiendo el orden de producción.
- v) Desarrollo de un gráfico de control para procesos binomiales con exceso de ceros o procesos en los que se han logrado estándares de calidad muy elevados.
- vi) Elaboración de programas en R para todas las funciones consideradas en la investigación.

### 7.3 Trabajos futuros

En este apartado se proponen algunas líneas para futuras investigaciones como continuación de los resultados obtenidos con este trabajo.

En el Capítulo 5, se efectuó la revisión de un conjunto de propuestas referidas a la implementación de gráficos de control clasificados en la categoría de gráficos de conformidades acumuladas. Se mostró la utilidad de emplear este tipo de gráficos para procesos de alta calidad particularmente procesos automatizados donde se puede establecer el orden de producción. Sin embargo, no fue posible ofrecer a los usuarios una metodología adecuada para el caso de procesos que se controlan por muestras. Esta es una investigación pendiente. Particularmente se podrían extender los resultados del procedimiento de optimización desarrollado por el Dr. J. T. Chen, para el caso secuencial (uno por vez), al caso en que la producción se controla por muestras.

Otra línea de investigación sugerida está referida a la posibilidad de extender los procedimientos de control para procesos de alta calidad al caso de procesos multinomiales. Los procesos multinomiales son procesos en los que cada unidad de producción es susceptible de ser clasificada, no sólo en dos categorías (conforme/no conforme), sino en una de varias categorías (bueno/regular/malo). Actualmente estos procedimientos de control existen sólo para muestras muy grandes, con proporción de no conformidades alta y altos volúmenes de producción.

Una línea de investigación altamente desafiante consistiría en establecer cómo se relacionan los procedimientos estadísticos propuestos para “mejorar” el control de características críticas de atributos, con los indicadores de capacidad del proceso. En 1983, el Dr. Calvin escribió: “Es importante reconocer que un camino hacia cero defectos implica una sinergia adecuada entre la fiabilidad de las especificaciones y la capacidad del proceso de fabricación. De particular preocupación son las técnicas estadísticas asociadas a los datos de atributos, donde se requiere una mayor conciencia de la relación entre las especificaciones y la capacidad del proceso de producción”. Por lo tanto sería muy importante estudiar cómo impacta la metodología estadística propuesta en esta investigación sobre los indicadores de capacidad o en su defecto, definir nuevos indicadores para procesos de atributos de alta calidad.

## Anexo I:

### I.1 Método de expansión de Cornish Fisher

La expansión de Cornish-Fisher es una fórmula para aproximar cuantiles de una variable aleatoria basada en sus cumulantes. Recordemos que los cumulantes de una variable aleatoria son similares a sus momentos y se definen como aquellos valores de  $K_r$  tal que la identidad:

$$\exp\left(\sum_{r=1}^{\infty} \frac{K_r t^r}{r!}\right) = \sum_{r=0}^{\infty} \frac{E(X^r) t^r}{r!} \quad \text{I.1}$$

se mantiene para todo  $t$ .

Los cumulantes de una variable aleatoria  $X$  pueden ser expresados en términos de su media,  $\mu = E(X)$  y momentos centrales  $\mu_r = E[(X - \mu)^r]$ .

La expresión para los primeros cinco cumulantes es:

$$K_1 = \mu, \quad K_2 = \mu_2, \quad K_3 = \mu_3, \quad K_4 = \mu_4 - 3\mu_2^2, \quad K_5 = \mu_5 - 10\mu_3\mu_2$$

Cornish y Fisher (1937) proveen una expansión para aproximar el cuantil  $q$ ,  $\Phi_X^{-1}(q)$  de una variable  $X$  normal con media 0 y desviación estándar unitaria, basada en sus primeros cumulantes, de la siguiente manera:

$$\begin{aligned} \Phi_X^{-1}(q) \approx & \Phi_Z^{-1}(q) + \frac{\Phi_Z^{-1}(q)^2 - 1}{6} K_3 + \frac{\Phi_Z^{-1}(q)^3 - 3\Phi_Z^{-1}(q)}{24} K_4 - \\ & - \frac{2\Phi_Z^{-1}(q)^3 - 5\Phi_Z^{-1}(q)}{36} K_3^2 + \frac{\Phi_Z^{-1}(q)^4 - 6\Phi_Z^{-1}(q)^2 + 3}{120} K_5 \\ & - \frac{\Phi_Z^{-1}(q)^4 - 5\Phi_Z^{-1}(q)^2 + 2}{24} K_3 K_4 + \frac{12\Phi_Z^{-1}(q)^4 - 53\Phi_Z^{-1}(q)^2 + 17}{324} K_3^3 \end{aligned} \quad \text{I.2}$$

Donde  $\Phi_Z^{-1}(q)$  es el cuantil  $q$  de una variable normal estándar

## I.2 La expansión de Cornish-Fisher aplicada al gráfico $p$

Para aplicar la expansión de Cornish-Fisher a los gráficos  $p$ , es necesario efectuar previamente el cálculo de los momentos y cumulantes de la distribución binomial.

### I.2.1 Momentos (naturales y centrados) de la distribución binomial

A partir de la función generadora de momentos de la distribución binomial, se obtienen los momentos naturales y centrados de la siguiente manera:

$$M_x(t) = \sum_{k=0}^n e^{tk} p_k = (pe^t + q)^n \quad \text{I.3}$$

donde  $p_k = \binom{n}{k} p^k q^{n-k}$ ,  $k = 0, 1, 2, \dots, n$  and  $q = 1 - p$

En particular, los primeros cuatro momentos naturales están dados por:

$$\begin{aligned} \mu'_1 &= \mathbb{E}(x) = np \\ \mu'_2 &= \mathbb{E}(x^2) = n^2 p^2 + np(1 - p) \\ \mu'_3 &= \mathbb{E}(x^3) = np(1 - p)(1 - 2p) + 3 n^2 p^2(1 - p) + n^3 p^3 \\ \mu'_4 &= \mathbb{E}(x^4) = np(1 - p)[1 - 6p(1 - p)] + 6 n^2 p^2(1 - p)[(1 - 2p) + \\ &\quad + np] + n^2 p^2(1 - p^2 + n^2 p^2) \end{aligned} \quad \text{I.4}$$

Los momentos centrales se obtienen a partir de la siguiente expresión:

$$\mu_h = E(X - \mu)^h = \sum_{r=0}^h \mu'_r \binom{h}{r} (-\mu)^{h-r} \quad h = 1, 2, \dots \quad \text{I.5}$$

$$\begin{aligned} \mu &= \mu'_1 = \mathbb{E}(X) = np \\ \mu_2 &= \mu'_2 - \mu_1'^2 = \mathbb{V}(X) = np(1 - p) \\ \mu_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3 = np(1 - p)(1 - 2p) \\ \mu_4 &= \mu'_4 - 4\mu'_1\mu'_3 + 6\mu_2'\mu_1'^2 - 3np(1 - p)[1 - 6p(1 - p)] + 3n^2 p^2(1 - p)^2 \end{aligned} \quad \text{I.6}$$

### I.2.2 Cumulantes

A partir de la función generadora de cumulantes (logaritmo de la función generadora de momentos) y, teniendo en cuenta los momentos previamente estandarizados, se obtiene:

$$\begin{aligned} K_1 &= \mu^* = \frac{\mu}{\sigma}, & K_2 &= \mu_2^* = \frac{\mu_2}{\sigma^2}, & K_3 &= \mu_3^* = \frac{\mu_3}{\sigma^3}, \\ K_4 &= \mu_4^* - 3\mu_2^{*2} = \frac{\mu_4}{\sigma^4} - 3\left(\frac{\mu_2}{\sigma^2}\right)^2 \end{aligned} \quad \text{I.7}$$

donde  $\sigma = \sqrt{\mu_2}$

## Anexo II:

### Publicaciones y Presentaciones en Congresos

En el marco de la presente investigación, se llevaron a cabo publicaciones y presentaciones en congresos de algunos de los contenidos de esta tesis. A continuación se detallan los trabajos más relevantes

#### II.1 Publicaciones

Con parte de los resultados obtenidos en los Capítulos 3 a 5 se realizaron los siguientes trabajos:

- 1.- Silvia Joekes y Emanuel Pimentel Barbosa (2013). “An improved attribute control chart for monitoring non-conforming proportion in high quality processes”. *Journal of Control Engineering Practice*, 21, 4, 407–412. ISSN: 0967-0661. Elsevier.
- 2.- Silvia Joekes, Emanuel Pimentel Barbosa, Marcelo Smrekar, (2014). “Extending a double sampling control chart for non-conforming proportion in high quality processes to the case of small samples”. *Journal of Statistical Methodology*, 23, 35–49. ISBN: 978-987-575. Elsevier.
- 3.- Silvia Joekes,, Marcelo Smrekar, Andrea Righetti, (2015). “A comparative study of two proposed CCC- $r$  charts for high quality processes and their application to an injection molding process". *Quality Engineering, Journal of the American Society for Quality* (Aprobado con revisión). ISSN: 0898-2112. Taylor & Francis

#### II.2 Presentaciones en Congresos

Además de las publicaciones detalladas anteriormente, con los resultados que fueron obteniéndose a lo largo de la investigación se realizaron las siguientes presentaciones en congresos nacionales y regionales:

1. S. Joekes, M. Smrekar, y E. Pimentel Barbosa. (2014). “Estudio del Gráfico de Control CCC- $r$  para Procesos de Alta Calidad y su Aplicación con Datos de una Planta de Autopartes”. XI Congreso Latinoamericano de Sociedades de Estadística (Clatse XI) y XLII Coloquio Argentino de Estadística, La Serena, Chile, 20/23 de Octubre. Publicado en el Libro de Resúmenes, p. 98.

2. A. Righetti, S. **Joekes**, y C. Abrego, (2014). “*Modelo zero-inflated Poisson en el control de procesos de alta calidad: un caso de aplicación*”. XI Congreso Latinoamericano de Sociedades de Estadística (Clatse XI) y XIII Coloquio Argentino de Estadística - La Serena, Chile, 20/23 de Octubre. Publicado en el Libro de Resúmenes, p. 127.
3. S. **Joekes**, M. Smrekar, y E. Pimentel Barbosa. (2013). “*Gráficos de control con muestreo simple y doble para la proporción de ítems no conformes en procesos de alta calidad*”. XLI Coloquio Argentino de Estadística, Facultad de Ciencias Económicas, Universidad Nacional de Cuyo, Mendoza, 16/18 de Octubre. Trabajo completo publicado en el Libro del Congreso, p.366-381. Publicado en el Libro de Resúmenes p. 28, ISBN: 978-987-575-116-3.
4. S. Buzzi , M. Llop , A. Righetti Y S. **Joekes**, (2013). “*Efecto de autocorrelación en gráficos c de control de atributos para procesos con exceso de ceros*”. XLI Coloquio Argentino de Estadística, Facultad de Ciencias Económicas, Universidad Nacional de Cuyo, Mendoza, 16/18 de Octubre. Publicado en el Libro de Resúmenes del Congreso, p. 127. Publicado en el Libro de Resúmenes, p. 65, ISBN: 978-987-575-116-3.
5. A. Righetti, S. **Joekes**, C. Abrego y M.Yacci, (2013). “*Modelos zero inflated Binomial y Poisson en el control estadístico de procesos de alta calidad*”. XLI Coloquio Argentino de Estadística, Facultad de Ciencias Económicas, Universidad Nacional de Cuyo, Mendoza, 16/18 de Octubre. Publicado en el Libro de Resúmenes, p. 66, ISBN: 978-987-575-116-3.
6. S. **Joekes**, M. Smrekar, y E. Pimentel Barbosa. (2012). “*Gráficos de control de atributos para procesos de alta calidad basados en el conteo acumulado de ítems conformes*”. X Congreso Latinoamericano de Sociedades de Estadística (CLATSE X) y XL Coloquio Argentino de Estadística, Universidad Nacional de Córdoba. Córdoba, 16/19 de Octubre. Publicado en el Libro de Resúmenes, ISBN 978-987-657-700-7.
7. A. Righetti, S. **Joekes**, M.Yacci y C. Abrego, (2012). “*Límites de control con ajustes en la expansión de Cornish-Fisher en gráficos U*. X Congreso Latinoamericano de Sociedades de Estadística (CLATSE X) y XL Coloquio Argentino de Estadística, Universidad Nacional de Córdoba. Córdoba, 16/19 de Octubre. Publicado en el Libro de resúmenes ISBN 978-987-657-700-7.
8. S. Buzzi , M. Llop , A. Righetti Y S. **Joekes**, (2012). “*Efectos de la autocorrelación en la gráfica c de control de control de atributos*”. X Congreso Latinoamericano de Sociedades de Estadística (CLATSE X) y XL Coloquio Argentino de Estadística, Universidad Nacional de Córdoba. Córdoba, 16/19 de Octubre. Trabajo publicado en el Libro de resúmenes ISBN 978-987-657-700-7.
9. S. **Joekes**, y E. Pimentel Barbosa. (2011) “*Propuesta de un gráfico p mejorado para procesos de alta calidad basado en la corrección cuantílica de Cornish-Fisher*”. XXXIX Coloquio Argentino de Estadística, Universidad Nacional del Litoral, Santa Fe, 12/14 de Octubre. Publicado en el Libro de resúmenes ISBN 978-987-657-700-7.

## Anexo III: Programas en R

### III.1 Programa para muestreo doble (Capítulo 4).

# Esta función calcula la probabilidad de ser necesaria una segunda muestra

```
SimulSM<-function(g,alfa0,p0,n,LSC,v1=2,v2=5){
```

```
#####
```

```
pac2<-function(LA,LSC1,LSC2,n1,n2,p){
```

```
pr<-0
```

```
for (i in (LA+1):(LSC1-1)){
```

```
  pr1<-dbinom(i,n1,p)
```

```
  pr2<-pbinom(LSC2-i,n2,p)
```

```
  a<-pr1*pr2
```

```
  pr<-pr+a
```

```
}
```

```
return(pr)
```

```
}
```

```
#####
```

# Esta rutina simula....

```
betainicial<-1
```

```
p1<-g*p0
```

```
for (n1 in seq(floor(0.5*n),ceiling(.8*n), by=v1)){
```

```
  for (LA in 0:LSC){
```

```
    for (LSC1 in (LA+2):(LA+LSC)){
```

```
      for (n2 in seq(n1,5*n1, by=v2)){
```

```
        LSC2max<-floor(.8*(LSC1*sqrt((n1+n2)/n1)))-1
```

```
        for(LSC2 in (LSC1+1):LSC2max){
```

```
          alfamin<-alfa0
```

```
          alfa_calc<-(1-(pbinom(LA,n1,p0)+pac2(LA,LSC1,LSC2,n1,n2,p0)))
```

```
          if (alfa_calc<=alfamin){
```

```
            alfamin<-alfa_calc
```

```
            nmin<-n
```

```
            n_calc<-n1+(n2*(dbinom(seq(LA+1,LSC1-1), n1, p0)%*%rep(1,LSC1-LA-1)))
```

```
            if (n_calc<= nmin){
```

```
              nmin<-n_calc;
```

```
              betamin<-betainicial
```

```
              beta_calc<-pbinom(LA, n1, p1)+pac2(LA,LSC1,LSC2,n1,n2,p1)
```

```
              if (beta_calc<betamin){
```

```
                betamin<-beta_calc
```

```
                if (beta_calc<betainicial){
```

```
                  betainicial<-beta_calc
```

```
                  n1ot<-n1
```

```
                  n2ot<-n2
```

```
                  LAot<-LA
```

```
                  LSC1ot<-LSC1
```

```
                  LSC2ot<-LSC2
```



```

}
}
}

return(c(LAot2,LSC1ot2,LSC2ot2,n1ot2,n2ot2,nmediot2,1/alfaat2,1/(1-betaot2)))
}

```

### III.2 Programa para gráficos tipo CCC (Capítulo 5)

**Referencias:** Chen0 = gráfico CCC de Xie et al. (2001), Chen1 = gráfico CCC-*r* de Chen y Cheng (2010), Chen2 = gráfico CCC-*r* de J. T. Chen (2009) y Zhang = gráfico CCS para muestras de Zhang (2005)

# Modificación de la función qcc del paquete qcc para que admita gráficos CCC-*r*

```

qccr<-function (data, type = c("xbar", "R", "S", "xbar.one", "p", "np",
" c", "u", "g", "gr"), sizes, center, std.dev, limits, data.name,
labels, newdata, newsizes, newlabels, nsigmas = 3, confidence.level,
rules = shewhart.rules, plot = TRUE, r=1,pp=0.1, ANI=0, Ch=0,...)
{
  call <- match.call()
  if (missing(data))
    stop("'data' argument is not specified")
  type <- match.arg(type)
  if (missing(data.name))
    data.name <- deparse(substitute(data))
  data <- data.matrix(data)

  if (missing(sizes)) {
    if (any(type == c("p", "np", "u")))
      stop(paste("sample 'sizes' must be given for a",
type, "Chart"))
    else sizes <- apply(data, 1, function(x) sum(!is.na(x)))
  }
  else {
    if (length(sizes) == 1)
      sizes <- rep(sizes, nrow(data))
    else if (length(sizes) != nrow(data))
      stop("sizes length doesn't match with data")
  }
  if (missing(labels)) {
    if (is.null(rownames(data)))
      labels <- 1:nrow(data)
    else labels <- rownames(data)
  }
  stats <- paste("stats.", type, sep = "")
  if (!exists(stats, mode = "function"))
    stop(paste("function", stats, "is not defined"))

#####

```

```

if (type=="gr"){
  if (missing(confidence.level)){aux<-0}
  else{aux<-confidence.level}
  stats<-do.call(stats, list(data=data, sizes=sizes, conf=aux, r=r, pp=pp, ANI=ANI, Ch=Ch))}
else {
  stats <- do.call(stats, list(data, sizes))}

#####

statistics <- stats$statistics
if (missing(center))
  center <- stats$center
sd <- paste("sd.", type, sep = "")
if (!exists(sd, mode = "function"))
  stop(paste("function", sd, "is not defined!"))
if (missing(std.dev)) {
  std.dev <- NULL
  std.dev <- switch(type, xbar = {
    if (any(sizes > 25)) "RMSDF" else "UWAVE-R"
  }, xbar.one = "MR", R = "UWAVE-R", S = "UWAVE-SD", NULL)
  std.dev <- do.call(sd, list(data, sizes, std.dev))
}
else {
  if (is.character(std.dev)) {
    std.dev <- do.call(sd, list(data, sizes, std.dev))
  }
  else {
    if (!is.numeric(std.dev))
      stop("if provided the argument 'std.dev' must be a method available or a numerical value.
See help(qcc).")
  }
}
names(statistics) <- rownames(data) <- labels
names(dimnames(data)) <- list("Group", "Samples")
object <- list(call = call, type = type, data.name = data.name,
  data = data, statistics = statistics, sizes = sizes,
  center = center, std.dev = std.dev)
if (!missing(newdata)) {
  newdata.name <- deparse(substitute(newdata))
  newdata <- data.matrix(newdata)
  if (missing(newsizes)) {
    if (any(type == c("p", "np", "u")))
      stop(paste("sample sizes must be given for a",
        type, "Chart"))
    else newsizes <- apply(newdata, 1, function(x) sum(!is.na(x)))
  }
  else {
    if (length(newsizes) == 1)
      newsizes <- rep(newsizes, nrow(newdata))
    else if (length(newsizes) != nrow(newdata))
      stop("newsizes length doesn't match with newdata")
  }
}
stats <- paste("stats.", type, sep = "")

```

```

if (!exists(stats, mode = "function"))
  stop(paste("function", stats, "is not defined"))

#####
if (type=="gr"){
  newstats<-do.call(stats,list(newdata, nwwsizes,confidence.level,r,pp,ANI,Ch))$statistics
}
else {
  newstats <- do.call(stats, list(newdata, newsizes))$statistics
}
#####

if (missing(newlabels)) {
  if (is.null(rownames(newdata))) {
    start <- length(statistics)
    newlabels <- seq(start + 1, start + length(newstats))
  }
  else {
    newlabels <- rownames(newdata)
  }
}
names(newstats) <- newlabels
object$newstats <- newstats
object$newdata <- newdata
object$newsizes <- newsizes
object$newdata.name <- newdata.name
statistics <- c(statistics, newstats)
sizes <- c(sizes, newsizes)
}
conf <- nsigmas
if (!missing(confidence.level))
  conf <- confidence.level
if (conf >= 1) {
  object$nsigmas <- conf
}
else if (conf > 0 & conf < 1) {
  object$confidence.level <- conf
}
if (missing(limits)) {
  limits <- paste("limits.", type, sep = "")
  if (!exists(limits, mode = "function"))
    stop(paste("function", limits, "is not defined"))

#####

if (type=="gr") {
  limits <- do.call(limits, list(center = center, std.dev = std.dev,
    sizes = sizes, conf = conf, r=r,pp=pp , ANI=ANI,Ch=Ch))
}
else {
#####

```

```

    limits <- do.call(limits, list(center = center, std.dev = std.dev,
                                sizes = sizes, conf = conf))
  }
  else {
    if (!missing(std.dev)) {
      warning("'std.dev' is not used when limits is given")
    }
    if (!is.numeric(limits))
      stop("'limits' must be a vector of length 2 or a 2-columns matrix")
    limits <- matrix(limits, ncol = 2)
    dimnames(limits) <- list(rep("", nrow(limits)), c("LCL ",
                                                       "UCL")) # los pone como lista, con nombre
  }
  lcl <- limits[, 1]
  ucl <- limits[, 2]
  object$limits <- limits
  if (is.function(rules))
    violations <- rules(object)
  else violations <- NULL
  object$violations <- violations
  class(object) <- "qcc"
  if (plot)
    plot(object, ...)
  return(object)
}

#####

limits.gr<-function (center, std.dev, sizes, conf , r , pp , ANI, Ch) {

  Chen1<-function(r,pp,conf){
    conf<-conf/2
    if (r==1)
      stop("For r=1 See Type=g.")
    else if (r>1 & r<5){
      b<-
matrix(c(1.1957779041,1.1241624759,1.090620321,1.0712911901,0.2587694457,0.1167280712,0
.0721220629,0.0514706254,0.0144695386,0.006048663,0.0032413867,0.0020023815), nrow = 4,
ncol = 3)
      k<-b[(r-1),1]-b[(r-1),2]*pp+b[(r-1),3]*log10(conf)
      a<-c(qnbinom(conf,r,pp)+r,qnbinom((1-conf),r,pp)+r)*k
    }
    else
      stop("invalid r argument. See help.")
    return(a)}

  Chen2<-function(r,pp,conf){
    slop<-function(r,pp,U,L){
      ar1lp<-(U*dnbinom(U-r, r, pp)-L*dnbinom(L-r, r, pp))/pp*(1-pnbinom(U-
r,r,pp)+pnbinom(L-r,r,pp))^2
      return(ar1lp)
    }
    meta<--Inf
    L<-r

```

```

fl<-0
while(fl<conf){
  fl1<-pnbinom(L-r+1,r,pp)
  stopu<-TRUE
  U<-qnbinom(1-conf,r,pp)+r-1
  while(stopu){
    fu<-pnbinom(U-r,r,pp)
    obj<-slop(r,pp,U,L)
    if ((fu-fl1)>=(1-conf)& (fu-fl1)<(1-conf)& obj<0){
      if(obj>meta){
        meta<-obj
        LI<-L
        LS<-U
      }
    }
    U<-U+1
    if(dnbinom(U-r,r,pp)<1e-5){stopu<-FALSE}
  }
  L=L+1
  fl<-fl1
}
return(c(LI,LS))}

Zhang<-function(n,r,pp,ANI){
  EMr<-function(n,r,pp){
    EM<-1/(1-(1-pp)^n)
    if(r>=2 & r<4){EM<-EM+(n*pp*(1-pp)^(n-1))/(1-(1-pp)^n)^2
      if(r==3){EM<-EM+(pp^2/2)*((n^2-n)*(1-pp)^(n-2)+(n^2+n)*(1-pp)^(2*n-2))/(1-(1-
pp)^n)^3}}
    else stop("invalid r argument. See help.")
    return(EM)}
  q<-EMr(n,r,pp)*n/ANI
  L<-(qnbinom(q,r,pp)+r)/n
  U<-(qnbinom(1-q,r,pp)+r)/n

  return(c(floor(L),floor(U))) }

if (missing(confidence.level) & ANI==0)
  stop("confidence.level or ANI must be given for a gr Chart")
else if (missing(confidence.level) & ANI!=0)
  limites <- Zhang(sizes[1],r,pp,ANI)
else if (confidence.level < 1 & confidence.level > 0) {
  if (ANI!=0)
    warning("'ANI' is not used when 'confidence.level' is given")
  if(Ch==0)
    limites <- c(qnbinom(conf,r,pp)+r,qnbinom((1-conf),r,pp)+r)
  else if (Ch==1)
    limites <- Chen1(r,pp,conf)
  else if (Ch==2)
    limites <- Chen2(r,pp,conf)
  else
    stop("invalid Ch argument. See help.")
}

```

```

    }
    else stop("invalid confidence.level argument. See help.")
    ucl <- limites[2]
    lcl <- limites[1]
    limites <- matrix(c(lcl, ucl), ncol = 2)
    rownames(limites) <- rep("", length = nrow(limites))
    colnames(limites) <- c("LCL", "UCL")
    return(limites)
}

#####

sd.gr<- function (data, sizes, ...){
  data <- as.vector(data)
  p <- 1/mean(data)
  std.dev <- sqrt((1 - p)/p)
  return(std.dev)
}

#####

stats.gr<-function (data=data, sizes = sizes, conf=conf , pp=pp, r=r, ANI=ANI,Ch=Ch) {
  statistics <- as.vector(data)
  if(Ch!=1) {
    center <- qnbinom(0.5,r,pp)+r}
  else{
    b<-
matrix(c(1.1957779041,1.1241624759,1.090620321,1.0712911901,0.2587694457,0.1167280712,0
.0721220629,0.0514706254,0.0144695386,0.006048663,0.0032413867,0.0020023815), nrow = 4,
ncol = 3)
    k<-b[(r-1),1]-b[(r-1),2]*pp+b[(r-1),3]*log10(conf)
    center <- (qnbinom(0.5,r,pp)+r)*k}
  list(statistics = statistics, center = center)
}

```

### III.3 Programa para datos de atributos con exceso de ceros (Capítulo 6)

(Requiere programas VEGAM, qcc y qcc-r)

```

inflados<-function(x,n){
  require(VGAM)
  sv<-rep(n,length(x))
  data<-cbind(x,sv)
  colnames(data)<-c("y", "z")
  fit1 <- vglm(cbind(x, sv - x) ~ 1, zibinomial, data = as.data.frame(data))
  fit2 <- vglm(cbind(x, sv - x) ~ 1, binomialff, data = as.data.frame(data))

  Estimaciones<-gl(3,1,6,labels=c("Akaike","fi","p"))
  Modelo<-gl(2,3,labels=c("Binomial","Zero-inflada"))
  y<-c(c(AIC(fit2),0,logit(coef(fit2),inverse=T),AIC(fit1)),logit(coef(fit1),inverse=T))
  tablex<-data.frame(y,Estimaciones,Modelo)

```

```
TablaDeSeleccion<-xtabs(y~Modelo+Estimaciones)
plot(table(x))
windows()
hist(x)
Salida<-list(summary(x),TablaDeSeleccion)
return(Salida)
}
```

---

## Referencias bibliográficas

- ✓ Acosta-Mejia, C. A. (1999). Improved p charts to monitor process quality. *IIE Transactions*, vol. 31, 509-516.
- ✓ Akaike, H. (1974), A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6): 716–723.
- ✓ Albers, W. (2010). The optimal choice of binomial negative charts for monitoring high quality processes. *Journal of Statistic Planning and Inference*. 140 (1), 214-225.
- ✓ Alwan, L. C. (2000). *Statistical Process Analysis*. Mc Graw-Hill
- ✓ Anscombe, F. J. (1948). The transformation of Poisson, binomial and negative binomial data. *Biometrika*, 35, 246–254.
- ✓ Bai, D.S. and Lee, K.T. (2002). Variable sampling interval X-bar control charts with an improved switching rule, *International Journal of Production Economics*, 54, 57.
- ✓ Bartlett, M. S. (1936). The square root transformation in analysis of variance. *The Journal of the Royal Statistical Society*, 3(1), 68–78.
- ✓ Bartlett, M. S. (1936). The use of transformations. *Biometrics*, 3(1), 39–52.
- ✓ Bourke, P.D., (1991). Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection. *Journal of Quality Technology*, 23, 225-238.
- ✓ Bourke, P.D., (2001). Sample size and the binomial CUSUM control chart: the case of 100% inspection. *Metrika* 53, 51-70.
- ✓ Calvin, T., (1983). Quality control techniques for zero defects. *IEEE Transactions on Components Hybrids and Manufacturing Technology*, 6(3), 323-328.
- ✓ Carot, V., Jabaloyes, J.M. and Carot, T. (2002). Combined double sampling and variable sampling interval X-bar chart. *International Journal of Production Research*, 40(9), 2175-2186.
- ✓ Cavalcanti, A. B. and Cordeiro, G. M. (2006). An improved *u* chart for attributes. *Brazil Journal of Probability and Statistic*, 20, 133-140.
- ✓ Chan, L.Y. y Wu, S. (2009). Optimal design for inspection and maintenance policy based on the CCC chart. *Computers & Industrial Engineering*, 57 (3), 667-676.
- ✓ Chan, L.Y., Lai, C.D., Xie, M. and Goh, T.N. (2003). A two-stage decision procedure for monitoring processes with low fraction nonconforming”. *European Journal of Operational Research*, 150, 420-436.
- ✓ Chan, L.Y., Lin, D.K.J., Xie, M. and Goh, T.N., (2002), Cumulative probability control charts for geometric and exponential process characteristics. *International Journal of Production Research* 38, 2, 397-408.
- ✓ Chan, L.Y., Xie, M. and Goh, T. N., (2000). Cumulative quantity control charts for monitoring production processes. *International Journal of Production. Research*, 38(2), 397-408.

- ✓ Chang, T.C. and Gan, F.F. (2001). Cumulative sum charts for high yield processes, *Statistica Sinica*, 11, 791-805.
- ✓ Chakraborti, S., Human, S.W y Graham, M.A. (2008). Phase I Statistical Process Control Charts: An Overview and some results. *Quality Engineering* 21:1, 52-62.
- ✓ Chen, G. (1998). An improved  $p$  chart through simple adjustment. *Journal of Quality Technology*, 30, 142-151.
- ✓ Chen, J. T. (2009). A new approach to setting control limits of cumulative count of conforming charts for high-yield processes. *Quality & Reliability Engineering International*, Vol. 25, No. 8, 973-986.
- ✓ Chen, G. y Cheng, S. W., (1998). The exact  $u$  chart obtained using simple adjustment. *Statistic & Probability Letters*, 37(4), 357-365.
- ✓ Chen, P. W and Cheng, CH. S. (2010). An ARL-unbiased Approach to Setting Control Limits of CCC- $r$  Chart for High-Yield Processes. *Journal of Quality*, Vol.17, N°6, 435-451.
- ✓ Chen, P.W. y Cheng, C.S. (2007). Cumulative count of conforming chart with variable sampling intervals for Markov dependent production processes. *Proceedings of the Second International Conference on Innovative Computing, Information and Control*.
- ✓ Chen, Y.K. (2004). Economic design of X-bar control charts for non-normal data using variable sampling policy. *International Journal of Production Economics*, 92(1), 61-74.
- ✓ Cornish EA, Fisher RA., (1937). Moments and cumulants in the specification of distributions. *Revue de l'Institut Internat de Statistic*; 4, 307–20.
- ✓ Cornish, E.A. and Fisher, R.A. (1960). “The Percentage Points of distributions Having Know Cumulants”. *Technometrics* 2, 209-225.
- ✓ Costa, A.F.B. and Claro, F.A.E., (2007). Double Sampling X Control Chart for a First-Order Autoregressive Moving Average Process Model. *International Journal of Advanced Manufacturing Technology*, on-line edition.
- ✓ Croasdale, R., (1974). Control Charts for a Double-Sampling Scheme Based On Average Production Run Lengths. *International Journal of Production Research*, 12(5), 585-592.
- ✓ Daudin, J.J., (1992). Double Sampling X Charts. *Journal of Quality Technology*, 24(2), 78-87.
- ✓ Deming, W.E; (1986). *Out of the Crisis*. MIT, Center for Advanced Engineering Study, Cambridge, MA.
- ✓ DeVor, R. E., Chang, T. and Sutherland, J. W. (2007). *Statistical Quality Design and Control*, 2<sup>nd</sup> edition, Pearson / Prentice-Hall.
- ✓ DeVor, R. E., Chang, T. and Sutherland, J. W. (1992). *Statistical Quality Design and Control*. Macmillan New York. York.
- ✓ Duncan, A.J. (1956). The economic design of X-bar charts used to maintain current control of a process, *Journal of the American Statistical Association*. 51, 228-242.

- 
- ✓ Fatahi, A.A; Noorosana, R; Dokouhaki, P. y Babakhani, M. (2010). Truncated Zero Inflated Binomial Control Charts for monitoring rare health events. *International Journal of Research and Reviews in Applied Sciences (IJRRAS)*, 4, 4, 380-387.
  - ✓ Freeman, M. F., & Tukey, J. W. (1950). Transformations related to the angular and the square root. *Annals of Mathematical Statistics*, 21, 607–611.
  - ✓ Goh T.N. and Xie M., (1995). Statistical process control for low nonconformity processes. *International Journal of Reliability, Quality and Safety Engineering*, 2, 15–22.
  - ✓ Goh, T. N., (1987). A control chart for very high yield processes. *Quality Assurance*, 13(1), 18-22.
  - ✓ He, D., Grigoryan, A. and Sigh, M., (2002). Design of Double- and Triple-Sampling X-bar Control Chart Using Genetic Algorithms. *International Journal of Production Research*, 40(6), 1387-1404.
  - ✓ Hinde, J.P.,& Demetrio, C.G.B. (1998). Overdispersion: models and estimation. *Computational Statistic and Data Analysis* 27, 151-170.
  - ✓ Hurvich, C.M. y Tsai, CH.L. (1989). Bias of the corrected AIC criterion for under fitted regression and time series models. *Biometrika* (1991), 78, 3, 499-509.
  - ✓ Irianto, D. and Shinozaki, N., (1998). An Optimal Double Sampling X Control Chart. *International Journal of Industrial Engineering: Theory, Applications and Practice*, 5(3), 226-234.
  - ✓ Kaminsky, F.C., Benneyan, J.C., Davis, R.D. and Burke, R.J. (1992). Statistical control charts based on a Geometric distribution. *Journal of Quality Technology*, Vol. 4, Nº 2, 63-69.
  - ✓ Katemee, N y Mayuresawan, T. (2012a). Nonconforming Control Charts for Zero Inflated Poisson distribution. *World Academy of Science, Engineering and Technology*. 69, 149-155.
  - ✓ Katemee, N y Mayuresawan, T. (2012b). Control Charts for Zero-Inflated Poisson Models. *Applied Mathematical Sciences*. 6-56, 2791-2803.
  - ✓ Kuralmani V., Xie M., Goh T.N. and Gan F.F., (2002). A conditional decision procedure for high yield processes. *IIE Transaction*, 34, 1021–1030.
  - ✓ Lai, C.D., Govindaraju, K. and Xie, M. (1998). Effects of correlation on fraction on conforming statistical process control procedures, *Journal of Applied Statistics*, 25(4), 535-543.
  - ✓ Lai, C.D., Xie, M. and Govindaraju, K. (2000). Study of a Markov model for a high quality dependent process, *Journal of Applied Statistics*, 27(4), 461-473.
  - ✓ Lee, K.T. and Bai, D.S. (2000). Variable sampling interval X-bar control charts with runs rules, *International Journal of Industrial Engineering: Theory, Application and Practice*, 7(2), 147-158.
  - ✓ Lee, Y. y Lee, M.C. (1992). On the derivation and Computation of the Cornish-Fisher Expansion. *Australian Journal of Statistic*. 34, 3, 443-450.
  - ✓ Lee, Y. S. and Lin, T. K. (1992). Higher-Order Cornish-Fisher Expansion. *Apply Statistics*, vol. 41, 233-240.

- 
- ✓ Lin, Y.C. and Chou, C.Y. (2005). On the design of variable sample size and sampling intervals X-bar charts under non-normality. *International Journal of Production Economics*, 26(2), 249-261.
  - ✓ Liu, J.Y., Xie, M. and Goh, T.N. (2006a). CUSUM chart with transformed exponential data, *Communications in Statistics: Theory and Methods*, 35, 1829-1843.
  - ✓ Lucas, J.M. (1989). Control scheme for low count levels. *Journal of Quality Technology*, 21, 199-201.
  - ✓ Madsen, R.W. (1993). Generalized binomial distributions, *Communications in Statistics: Theory and Method*, 22, 3065-3086.
  - ✓ McCullagh & Nelder (1989) *Generalized Linear Models*. Chapman and Hall.
  - ✓ Montgomery, D .C. (2008). *Introduction to Statistical Quality Control*. (6th ed.), John Wiley & Sons.
  - ✓ Montgomery, D. C. (1991). *Introduction statistical quality control*. (1st ed.), John Wiley & Sons.
  - ✓ Nelder J. A. and Wedderburn R. W. M (1972). "Generalized Lineal Model", *Journal of the Royal Statistical Society. Series A (General)*, Vol. 135, No. 3, 370-384.
  - ✓ Nelson, L.S. (1994). A control chart for parts-per-million nonconforming items. *Journal of Quality Technology*, 26, 239-240.
  - ✓ Ohta, H., Kusukawa, E. and Rahim, A. (2001). A CCC-r chart for high-yield processes. *Quality and Reliability Engineering International*, 17, 439-446.
  - ✓ Peerajit V., y Mayuresawan T. (2010). Nonconforming Control Charts for Zero-inflated processes. *Proceedings of the National Conference on Statistics and Applied Statistics*, 61-73.
  - ✓ Quesenberry, C. P. (1991). SPC Q Charts for Binomial Parameter  $p$ : Short or Long Runs. *Journal of Quality Technology* 23, 239-246.
  - ✓ Quesenberry, C. P. (1995). On Properties of Binomial Q Charts for Attributes. *Journal of Quality Technology*, 27, 204-213.
  - ✓ Rodrigues, A.A, (2005) "Gráficos de controle por atributos com amostragem dupla" Tesis de doutorado. Rio de Janeiro- Brasil.
  - ✓ Rodrigues, A. A., Epprecht, E.,K., Magalhães, M.S. (2011). Double-sampling control charts for attributes. *Journal of Applied Statistics-*, Vol. 38, N° 1, 87-112.
  - ✓ Ryan T.P., and Schwertman, N.C, (1997). Optimal limits for attributes control charts. *Journal of Quality Technology*, 29, 86-98.
  - ✓ Ryan, T. P. (1989). *Statistical Methods for Quality Improvement*. John Wiley & Sons
  - ✓ Chakraborti, S., Human, S.W. y Graham, M.A., (2008). Phase I Statistical Process Control Charts: An Overview and some results. *Quality Engineering*, 21:1, 52-62.
  - ✓ Shewhart, W.A., (1931). Economic Control of Quality of Manufactured Product. *Journal of the American Statistical Association*. Vol. 27, N° 178, 215-217.
  - ✓ Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* 6 (2): 461-464.

- 
- ✓ Shewhart, W.A., (1926). Quality Control Charts. Bell System. *Technical Journal*, 5, 593-603.
  - ✓ Sim, C.H. y Lim, M.H., (2008). Attribute Charts for Zero-Inflated Processes. *Communications in Statistics, Simulation and Computation*, 37: 1440-1452.
  - ✓ Surath Aebtarm y Nizar Bougila, (2011). An empirical evaluation of attribute control charts for monitoring defects. *Journal of Expert Systems with Applications*. Volume 38 Issue 6, 7869-7880
  - ✓ Surucu, B. and Sazak, H.S. (2009). Monitoring reliability for a three-parameter Weibull distribution. *Reliability Engineering and System Safety*, 94, 503-508.
  - ✓ Tang, L.C. and Cheong, W.T. (2006). A control scheme for high-yield correlated production under group inspection. *Journal of Quality Technology*, 38(1), 45-55.
  - ✓ Tsai, T., Lin, C. y Wu, S. (2006). Alternative Attribute Control Charts based on Improved Square Root Transformation. *Journal of Mathematic Sciences* 22(1), 61-72.
  - ✓ Vieira, A. M. C., Hinde, J. P., Demetrio, C. G. B. (2000). Zero-inflated proportion data models applied to a biological control assay. *Journal of Applied Statistics* 27:373-389.
  - ✓ Villalobos, J.R., Munoz, L. y Gutierrez, M.A. (2005). Using fixed and adaptive multivariate SPC charts for online SMD assembly monitoring. *International Journal of Production Economics*. 95(1), 109-121.
  - ✓ Wang, H. (2009). Comparison of  $p$  control charts for low defective rate. *Computational Statistics and Data Analysis*. 53, 12, 4210-4220.
  - ✓ Winterbottom, A. (1993). Simple adjustments to improve control limits on attribute charts. *Quality and Reliability Engineering International*, 9, 2, 105-109.
  - ✓ Woodall, W. H. (1997). Control charts based on attribute data. *Journal of Quality Technology*, 29, 172-183.
  - ✓ Wu Z. and Wang, Q. (2007). An np control chart using double inspections. *Journal of Apply Statistic* 34(7), 843-855.
  - ✓ Wu, Z. and Luo, H. (2004). Optimal design of the adaptive sample size and sampling interval np control chart. *Quality and Reliability Engineering International*, 20(6), 553-570.
  - ✓ Wu, Z. and Spedding, T.A. (1999). Evaluation of ATS for CRL control chart. *Process Control and Quality*, 11(3), 183-191.
  - ✓ Wu, Z., Luo, H. y Zhang, X. (2006). Optimal np control chart with curtailment. *European Journal of Operational Research*, 174, 1723-1741.
  - ✓ Wu, Z., Yeo, S.H. and Spedding, T. (2001). A synthetic control chart for detecting fraction nonconforming increases. *Journal of Quality Technology*, 33(1), 104-111.
  - ✓ Wu, Z., Zhang, X., Yeo, S.H., (2001). Design of the sum-of-conforming-run-length control charts. *European Journal of Operational Research*, 132, 187-196.
  - ✓ Xie M. and Goh T.N., (1992). Some procedures for decision making in controlling high yield processes. *Quality and Reliability Engineering International* 8, 355-360.

- 
- ✓ Xie M. and Goh T.N., (1993). Improvement detection by control charts for high yield processes. *International Journal of Quality and Reliability Management* 10, 24–31.
  - ✓ Xie, M., Goh, T. N., and Kuralmani, V. (2000). On optimal setting of control limits for geometric chart. *International Journal of Reliability Quality and Safety Engineering*, 7, 1, 17-25.
  - ✓ Xie, M., Goh, T.N. and Ranjan, P. (2002b). Some effective control chart procedures for reliability monitoring. *Reliability Engineering and System Safety*, 77, 143-150.
  - ✓ Xie, M., He, B., Goh, T. N. (2001). Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis* 38:191–201.
  - ✓ Xie, M., Lu, X. S., Goh, T. N. and Chan, L. Y., (1999). A quality monitoring and decision-making scheme for automated production processes. *International Journal of Quality and Reliability Management*, 16(2), 148-157.
  - ✓ Xie, M.; Goh, T. N. and Kuralmani,V. (2002). *Statistical Models and Control Charts for High Quality Processes*. Boston, Kluwer Academic Publisher.
  - ✓ Yawsaeng, B. y Mayuresawan, T. (2012). Control Charts for Zero-Inflated Binomial Models. *Thailand Statistician*, 10(1), 107-120.
  - ✓ Yeh, A. B., McGrath, R. N., Sembower, M. A., and Shen, Q., (2008). EWMA control charts for monitoring high yield processes based on non-transformed observations. *International journal of Production Research*, 46(20), 5679-5699.
  - ✓ Yip, P. (1988). Inference about the mean of a Poisson distribution in the presence of a nuisance parameter. *Australian Journal of Statistics* 30:299–306.
  - ✓ Zhang L, Govindaraju K, Bebbington M, Lai CD. (2004). On the statistical design of geometric control charts. *Quality Technology and Quantitative Management*; 1(2): 233–243.
  - ✓ Zhang, C.W., Xie, M. and Goh, T.N. (2005). Economic design of exponential charts for time between events monitoring. *International Journal of Production Research*, Vol. 43, 23(1), 5019-5032.
  - ✓ Zhang, C.W., Xie, M. and Goh, T.N. (2006). Design of exponential control charts using a sequential sampling scheme. *IIE Transactions*, 38, 1105-1116.
  - ✓ Zhang, C.W., Xie, M. and Goh, T.N. (2008). Economic design of cumulative count of conforming charts under inspection by samples. *International Journal of Production Economics*, 111, 93-104.
  - ✓ Zhang, C.W., Xie, M., Liu, J.Y. and Goh, T.N. (2007). A control chart for the Gamma distribution as a model of time between events. *International Journal of Production Research*, 45(23), 5649-5666.
  - ✓ Zhang, C.W., Xie, M., Jin, T. (2012). An improved self-starting cumulative count of conforming chart for monitoring high-quality processes under group inspection. *International Journal of Production Research*, 1-18.