



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Departamento de Estadística e Investigación
Operativa Aplicadas y Calidad

Doctorado en Estadística y Optimización

**Gráficos de control por atributos con
curvas ARL cuasi insesgadas: Análisis y
desarrollo de métodos**

TESIS DOCTORAL

Presentada por:

Marco Antonio Argoti Morales

Dirigida por:

Dr. D. Andrés Carrión García

Valencia – España

Febrero de 2019

Dedicatoria:

A toda mi familia, tanto del viejo como
del nuevo continente, a mi hijo y hermanos.

En especial a mi madre, Carmen Amelia.

Agradecimientos:

Deseo agradecer a todas aquellas personas e instituciones que han contribuido directa e indirectamente a la realización de esta tesis doctoral.

Primeramente debo agradecer de manera muy especial a mi tutor y director de tesis Dr. Andrés Carrión García, por su guía, apoyo y sabios consejos.

A mis hermanos y familia, quienes siempre me han respaldado incondicionalmente para lograr y superar retos académicos y de vida.

A mi hermana, Sonia Angélica, quien ha sido el pilar en el cual me he sostenido para poder alejarme de casa durante la duración de mis estudios de postgrado.

A toda la gente y amigos que en estos años he conocido en Valencia, quienes, de una forma u otra, me han apoyado haciendo que mi estancia en esta maravillosa ciudad haya sido de lo más agradable.

A la Universidad Politécnica de Valencia y a su Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, por la formación recibida en estos años.

A la Secretaria de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) del Ecuador, por auspiciar y darme la oportunidad de realizar mis estudios doctorales en España.

A todos gracias.

Resumen:

Los gráficos de control por atributos son herramientas estadísticas ampliamente utilizadas tanto en la industria de bienes como en la de servicios y sirven para monitorizar procesos mediante atributos de calidad. Los gráficos por atributos uni-variantes están entre los más conocidos y son el tema central de esta tesis. Estos gráficos se dividen en dos tipos: los basados en la distribución binomial (Gráficos np y p) y aquellos basados en la distribución de Poisson (Gráficos c y u).

Los gráficos por atributos Shewhart son sin lugar a duda los más populares y tienen la particularidad de que sus límites de control se basan en la aproximación normal a las distribuciones binomial y de Poisson. Cuando se utilizan estos gráficos lo habitual es asumir que, siempre y cuando la aproximación normal sea adecuada, su capacidad de monitorización será idónea, es decir que podrán detectar de igual manera tanto mejoras como deterioros del proceso.

En esta tesis se demuestra que, debido a la asimetría de las distribuciones binomial y de Poisson, el ajuste de la aproximación normal es impreciso en las colas de esas distribuciones y que eso afecta negativamente la potencia de detección de los gráficos Shewhart. Para poder establecer la magnitud de la afectación, se desarrollaron varios parámetros novedosos que sirven para evaluar y caracterizar la capacidad de monitorización de cualquier gráfico por atributos del tipo uni-variante. Por medio de estos parámetros se estableció que los gráficos Shewhart, al contrario de lo que se presume, están lejos de ser idóneos.

Los nuevos parámetros mencionados en el párrafo anterior, también sirvieron para analizar gráficos de control planteados como alternativas superiores a los Shewhart. Los resultados de los análisis demostraron que esos gráficos tampoco tienen una capacidad de monitorización del todo satisfactoria.

Dos nuevos gráficos de control, el p Kmod y el u Kmod, son propuestos. Estos gráficos tienen una capacidad de monitorización superior a cualquier otro gráfico de control (p y u respectivamente) incluido en esta tesis y además cuentan con un método de fácil uso mediante el cual es posible establecer si esa capacidad es o no óptima.

Los resultados de la investigación han sido publicados en actas de congresos y en revistas científicas internacionales.

Resum:

Els gràfics de control per atributs són ferramentes estadístiques àmpliament utilitzades tant en la indústria de béns com en la de serveis i serveixen per a monitoritzar processos per mitjà d'atributs de qualitat. Els gràfics per atributs uni- variants estan entre els més coneguts i són el tema central d'esta tesi. Existeixen dos tipus: els gràfics basats en la distribució binomial (Gràfics np i p) i els gràfics basats en la distribució de Poisson (Gràfics c i u).

Els gràfics per atributs Shewhart són sens dubte els més populars i tenen la particularitat que els seus límits de control es basen en l'aproximació normal a les distribucions binomial i de Poisson. Quan s'utilitzen estos gràfics allò més habitual és assumir que, sempre que l'aproximació normal siga adequada, la seua capacitat de monitorització serà idònia, és a dir que podran detectar de la mateixa manera tant millores com deterioraments del procés.

En aquesta tesi es demostra que, a causa de la asimetria de les distribucions binomial i de Poisson, l'ajust de l'aproximació normal és imprecís en les cues d'eixes distribucions i que això afecta negativament la potència de detecció dels gràfics Shewhart. Per a poder establir la magnitud de l'afectació es van desenvolupar diversos paràmetres nous que servixen per a avaluar i caracteritzar la capacitat de monitorització de qualsevol gràfic per atributs del tipus univariant. A través d'ells es va establir que els gràfics Shewhart, al contrari del que es presumeix, estan lluny de ser idonis.

Els nous paràmetres mencionats en el paràgraf anterior també van servir per a analitzar gràfics de control plantejats com a alternatives superiors als Shewhart. Els resultats de les anàlisis van demostrar que tampoc tenen una capacitat de monitorització del tot satisfactòria.

Dos nous gràfics de control, el p K_{mod} i el u K_{mod} , són proposats. Estos gràfics tenen una capacitat de monitorització superior a qualsevol altre gràfic de control (p i u respectivament) inclòs en esta tesi, a més de comptar amb un mètode de fàcil ús, per mitjà del qual és possible establir si eixa capacitat és òptima o no.

Els resultats de la investigació han sigut publicats en actes de congressos i en revistes científiques internacionals.

Abstract:

Attribute control charts are statistical tools that are widely used in the goods and services industries, they serve to monitor processes by means of product quality attributes. The univariate attribute charts are amongst the most well-known and are the main topic of this thesis. Two types of such charts exist, namely: the ones based on the binomial distribution (np and p Charts) and the ones based on the Poisson distribution (c and u Charts).

The Shewhart attribute charts are without doubt the most popular and have the peculiarity that their control limits are based on the normal approximation to the binomial and Poisson distributions. When these charts are used it is commonly assumed that, as long as the normal approximation is adequate, their monitoring capability will be ideal, or in other words, that they will be able to detect with equal capacity, either process improvements or deteriorations.

In this thesis we show that due to asymmetry of the binomial and Poisson distributions, the adjustment of the normal approximation is inaccurate on their tail sides and that this affects on a detrimental way the detection power of the Shewhart charts. In order to be able to establish the magnitude of the affectation, various novel parameters that serve to assess and characterised the monitoring capability of any uni-attribute type chart were developed. Through them it was established that the Shewhart charts, contrary to what is commonly assumed, are far from being ideal.

The aforementioned novel parameters, also served to analyse other control charts posed as superior alternatives to the Shewhart's. The analysis results demonstrated that those charts, although superior to the Shewhart's, also have a far from satisfactory monitoring capability.

Two new control charts, the p Kmod and the u Kmod, are proposed. These charts have a superior monitoring capability compared to any other chart (p and u respectively) included in this thesis, in addition they have an easy to use method that makes it possible to establish if their monitoring capability is, or is not, ideal.

The results of the research have been published in congress proceedings and international scientific journals.

1 Contenido General

1.1 Índice de contenidos

1	CONTENIDO GENERAL.....	I
1.1	ÍNDICE DE CONTENIDOS	I
1.2	ÍNDICE DE TABLAS	VI
1.3	ÍNDICE DE FIGURAS	VIII
1.4	ÍNDICE DE ECUACIONES.....	XII
1.5	ACRÓNIMOS.....	XIV
1.6	NOMENCLATURA	XV
2	INTRODUCCIÓN Y OBJETIVOS.....	1
2.1	INTRODUCCIÓN	1
2.2	PLANTEAMIENTO DEL PROBLEMA	2
2.3	OBJETIVOS DE LA INVESTIGACIÓN	3
2.3.1	<i>Objetivo general</i>	3
2.3.2	<i>Objetivos específicos</i>	3
2.4	ESTRUCTURA DE LA TESIS	4
3	ANTECEDENTES RESPECTO AL TEMA DE INVESTIGACIÓN	6
3.1	EL ROL DE LOS GRÁFICOS DE CONTROL EN LA OBTENCIÓN DE LA CALIDAD	6
3.2	GRÁFICOS DE CONTROL	8
3.2.1	<i>Generalidades</i>	8
3.2.2	<i>Tipos de gráficos de control</i>	12
3.2.2.1	Gráficos de control uni-variantes basados en variables continuas	13
3.2.2.2	Gráficos de control uni-variantes basados en atributos de calidad	13
4	MARCO TEÓRICO BASE.....	16
4.1	GRÁFICOS DE CONTROL POR ATRIBUTOS BASADOS EN LA DISTRIBUCIÓN BINOMIAL	16
4.1.1	<i>Introducción al uso de la distribución binomial en gráficos de control por atributos</i>	16
4.1.2	<i>Propiedades de la distribución binomial $B(n, p)$</i>	17
4.1.2.1	Media y Varianza de la distribución binomial mediante expansión binomial.....	17
4.1.2.2	Media y Varianza de $B(n, p)$ mediante generación de momentos	19
4.1.2.3	Simetría de la distribución binomial	20
4.1.2.4	Sinopsis de la sección 4.1.2.....	24
4.1.3	<i>Aproximación de la distribución binomial a la normal</i>	25
4.1.3.1	Características de la distribución normal	25

4.1.3.2	Comprobación de la aproximación normal a la distribución binomial	25
4.1.4	<i>Principios de gráficos de control np y p</i>	28
4.1.4.1	Ecuaciones para computar los límites de control de los gráficos de control np y p	30
4.1.4.2	Relación entre prueba de hipótesis y los gráficos de control por atributos	31
4.1.4.3	Variaciones del proceso y como se manifiestan en el gráfico de control	33
4.1.5	<i>El ARL: Un parámetro para evaluar la capacidad de los gráficos de control por atributos para detectar cambios en el proceso</i>	35
4.1.5.1	ARL para el gráfico de control p cuando el proceso está bajo control (ARL ₀)	36
4.1.5.2	ARL para el gráfico de control p cuando el proceso está fuera de control (ARL ₁)	37
4.1.6	<i>Gráfico de Control p Shewhart: Colas alfa y ARL₀</i>	38
4.1.6.1	Comportamiento real de las colas alfa para el gráfico de control p Shewhart	39
4.1.6.2	Comportamiento real del ARL ₀ para el gráfico de control p Shewhart	41
4.1.7	<i>Análisis del ajuste de la aproximación normal a la distribución binomial cuando p << 0.5.....</i>	42
4.1.7.1	El ratio de las colas alfa.....	45
4.1.8	<i>Gráfico de control p Shewhart : El ARL₁ y la curva ARL.....</i>	45
4.1.8.1	Parámetros para determinar la severidad del sesgo en la curva ARL	47
4.1.8.2	Gráfico de control p Shewhart: variación de la severidad del sesgo en las curvas ARL.....	48
4.1.9	<i>Sinopsis de la sección 4.1.....</i>	49
4.2	GRÁFICOS DE CONTROL POR ATRIBUTOS BASADOS EN LA DISTRIBUCIÓN DE POISSON	50
4.2.1	<i>Introducción al uso de la distribución de Poisson en gráficos de control por atributos.....</i>	50
4.2.2	<i>Propiedades de la distribución de Poisson - Poi(c).....</i>	51
4.2.3	<i>Aproximación normal a la distribución de Poisson.....</i>	52
4.2.3.1	Ajuste de la aproximación normal a la distribución de Poisson.....	53
4.2.4	<i>Principios de gráficos de control c y u Shewhart.....</i>	55
4.2.4.1	ARL para el gráfico de control u cuando el proceso está bajo control (ARL ₀)	57
4.2.4.2	ARL para el gráfico de control u cuando el proceso está fuera de control (ARL ₁)	58
4.2.5	<i>El ARL₀ en gráficos de control u Shewhart.....</i>	58
4.2.5.1	Comportamiento real de las colas alfa para el gráfico de control u Shewhart	59
4.2.5.2	Comportamiento real del ARL ₀ en el gráfico de control u Shewhart.....	61
4.2.6	<i>La curva ARL para el gráfico de control u Shewhart.....</i>	62
4.2.6.1	Gráfico de control u Shewhart: Severidad del sesgo en las curvas ARL.....	63
4.2.7	<i>Sinopsis de la sección 4.2.....</i>	63
5	RESULTADOS Y CONTRIBUCIONES AL CONOCIMIENTO	64
5.1	CARACTERIZACIÓN Y ANÁLISIS DE LAS COLAS ALFA Y EL ARL ₀	64
5.1.1	<i>Método gráfico para analizar el comportamiento de las colas alfa.....</i>	64
5.1.2	<i>Método gráfico para analizar el comportamiento del ARL₀.....</i>	65
5.1.3	<i>Determinación del rango de ARL₀ aceptables</i>	65
5.1.4	<i>El ratio de las colas alfa (R_α).....</i>	67
5.2	GRÁFICO DE CONTROL P SHEWHART: CARACTERIZACIÓN DEL ARL-SESGO(%) Y R _A EN FUNCIÓN DE P Y N	67

5.3	GRÁFICOS DE CONTROL P SHEWHART: NUEVOS MÉTODOS PARA ESTIMAR LA SEVERIDAD DEL SESGO DE LAS CURVAS ARL	70
5.3.1	<i>Ecuación para estimar la severidad del sesgo en la curva ARL</i>	70
5.3.1.1	Curvas ARL cuasi insesgadas: definición	74
5.3.2	<i>Zonas de severidad del sesgo para curvas ARL</i>	74
5.4	EL ARL_{BSL} (ARL BIAS SEVERITY LEVEL) O NIVEL DE SEVERIDAD DEL SESGO EN LA CURVA ARL	77
5.4.1	<i>El parámetro ARL_{BSL}: Definición y criterios para interpretarlo</i>	77
5.4.2	<i>Criterio ARL_{BSL} para obtener curvas ARL cuasi insesgadas</i>	78
5.5	EL GRÁFICO DE CONTROL P KMOD: UN NUEVO GRÁFICO PARA MONITORIZAR PROCESOS CON BASE EN LA DISTRIBUCIÓN BINOMIAL	79
5.5.1	<i>Gráfico de control p Shewhart: Caracterización del ARL_{BSL}</i>	79
5.5.1.1	GC p Shewhart: relación entre las colas alfa y el ARL_{BSL}	81
5.5.1.2	Sinopsis de la sección 5.5.1	81
5.5.2	<i>Gráficos de Control p alternativos: Caracterización del ARL_0 y ARL_{BSL}</i>	82
5.5.2.1	Gráficos de control p : Enfoque optimización de los límites de control	83
5.5.2.1.1	Gráfico de control p basado en la expansión Corner-Fisher	83
5.5.2.1.2	Gráfico de Control p R&S - basado en análisis de regresión	87
5.5.2.2	Gráficos de control p : Enfoque uso de transformaciones no lineales	89
5.5.2.2.1	Gráfico de Control p basado en la transformación Arco Seno	90
5.5.2.2.2	Gráfico de Control p basado en la transformación raíz cuadrada	93
5.5.2.2.3	Gráfico de control p basado en la transformación Q	94
5.5.2.2.4	Comparativa de los resultados ARL_{BSL} para los gráficos p basados en transformaciones	94
5.5.2.3	Sinopsis de la sección 5.5.2	95
5.5.3	<i>El gráfico de control p Kmod</i>	96
5.5.3.1	El factor K adaptativo y los límites de control para el gráfico de control p Kmod	96
5.5.3.2	Gráfico de control p Kmod: Caracterización del ARL_0 y ARL_{BSL}	98
5.5.3.3	Gráfico de control p Kmod: Comparación con otros gráficos de control	100
5.5.3.4	Gráfico de control p Kmod: Método para determinar si una curva ARL es cuasi insesgada	102
5.5.3.5	Gráfico de control p Kmod: ejemplo de curvas ARL	103
5.5.3.6	Gráfico de control p Kmod: Aplicación al caso de monitorización de procesos mediante tamaños de muestra variables	105
5.6	EL GRÁFICO DE CONTROL U KMOD: UN NUEVO GRÁFICO PARA MONITORIZAR PROCESOS CON BASE EN LA DISTRIBUCIÓN DE POISSON	106
5.6.1	<i>Gráfico de control u Shewhart: Caracterización del ARL_{BSL}</i>	106
5.6.1.1	Gráfico de control u Shewhart: relación entre las colas alfa y el ARL_{BSL}	108
5.6.1.2	Sinopsis de la sección 5.6.1	109
5.6.2	<i>Gráficos de control u alternativos: Caracterización del ARL_0 y ARL_{BSL}</i>	109
5.6.2.1	Gráfico de control u basado en la expansión Corner-Fisher	110
5.6.2.2	Gráfico de control u basado en análisis de regresión	113
5.6.2.3	Gráfico de control u basado en la transformación "Power Transformación"	115
5.6.2.4	Sinopsis de la sección 5.6.2	118
5.6.3	<i>El gráfico de control u Kmod</i>	118

5.6.3.1	El factor K adaptativo y los límites de control para el gráfico de control u Kmod.....	118
5.6.3.2	Gráfico de control u Kmod: Caracterización del ARL_{BSL} y ARL_0	120
5.6.3.3	Gráfico de control u Kmod: Comparación con otros gráficos de control	122
5.6.3.4	Gráfico de control u Kmod: Método para determinar si una curva ARL es cuasi insesgada	123
5.6.3.5	Gráfico de control u Kmod: ejemplo de curvas ARL.....	124
5.6.3.6	Gráfico de control u Kmod: Aplicación al caso de monitorización de procesos mediante tamaños de muestra variables.....	126
5.7	APLICACIÓN DE UN GRÁFICO DE CONTROL KMOD A UN CASO REAL	126
6	CONCLUSIONES GENERALES, PRINCIPALES APORTACIONES Y POSIBLES TEMAS DE INVESTIGACIÓN...	129
6.1	CONCLUSIONES GENERALES	129
6.1.1	<i>Conclusiones concernientes al objetivo específico número uno.....</i>	<i>129</i>
6.1.2	<i>Conclusiones concernientes al objetivo específico número dos</i>	<i>130</i>
6.1.3	<i>Conclusiones concernientes al objetivo específico número tres.....</i>	<i>131</i>
6.1.4	<i>Conclusiones concernientes al objetivo específico número cuatro.....</i>	<i>131</i>
6.1.5	<i>Conclusiones concernientes al objetivo específico número cinco.....</i>	<i>132</i>
6.2	PRINCIPALES APORTACIONES OBTENIDAS DE LA INVESTIGACIÓN	133
6.3	POSIBLES TEMAS PARA FUTURA INVESTIGACIÓN	134
7	REFERENCIAS	135
8	ANEXOS.....	140
8.1	ANEXO 1	140
8.1.1	<i>Gráfico de control p Shewhart: Comportamiento de las colas alfa y el ARL_0 (ver sección 4.1.6)</i>	<i>140</i>
8.2	ANEXO 2	142
8.2.1	<i>Ejemplo de ajuste de la aproximación normal a la binomial para $n = 600$ y $p = 0.05$ (ver sección 4.1.7)</i>	<i>142</i>
8.3	ANEXO 3	143
8.3.1	<i>Gráfico de control p Shewhart: Ejemplos de curvas ARL (ver sección 4.1.8)</i>	<i>143</i>
8.4	ANEXO 4	144
8.4.1	<i>Gráfico de control p Shewhart: Caracterización del ARL-sesgo(%) y R_α (ver sección 5.2)</i>	<i>144</i>
8.5	ANEXO 5	146
8.5.1	<i>Gráfico de control p Shewhart: Ejemplos de resultados ARL_{BSL} (ver sección 5.5.1)</i>	<i>146</i>
8.6	ANEXO 6	147
8.6.1	<i>Gráfico de control p Kmod: Ejemplos de resultados ARL_{BSL} y ARL (ver sección 5.5.3.2)</i>	<i>147</i>
8.7	ANEXO 7	149
8.7.1	<i>Gráfico de control u Shewhart: Comportamiento de las colas alfa y el ARL_0 (ver secciones 4.2.5.1 y 4.2.5.2)</i>	<i>149</i>
8.8	ANEXO 8	151
8.8.1	<i>Gráfico de control u Shewhart: Ejemplos de Curvas ARL (ver sección 4.2.6).....</i>	<i>151</i>

8.9	ANEXO 9	152
8.9.1	<i>Gráfico de control u Shewhart: Ejemplos de resultados ARL_{BSL} (ver sección 5.6.1)</i>	<i>152</i>
8.10	ANEXO 10	153
8.10.1	<i>Gráfico de control u Kmod: Ejemplos de resultados ARL_0 y ARL_{BSL} (ver sección 5.6.3.2)</i>	<i>153</i>
8.11	ANEXO 11	154
8.11.1	<i>Algunos códigos R utilizados en la investigación</i>	<i>154</i>

1.2 Índice de tablas

Tabla 1: Las ocho dimensiones de la calidad de un producto según Garvin	6
Tabla 2: Ejemplo de resultados de muestreo para determinar p , utilizando $n = 100$	29
Tabla 3: Ecuaciones para calcular los límites de control para los gráficos np y p tipo Shewhart para un tamaño de muestra n	30
Tabla 4: Relación entre prueba de hipótesis y los GCs por atributos.....	31
Tabla 5: GC p Shewhart – ejemplo de valores ARL_1	46
Tabla 6: Definición de los parámetros de la curva ARL	47
Tabla 7: Ecuaciones para calcular los límites de control para los gráficos c y u tipo Shewhart	56
Tabla 8: Criterios para determinar la severidad del sesgo en función del ARL-sesgo(%).....	73
Tabla 9. Criterios para establecer los límites de las zonas de severidad	75
Tabla 10. Ejemplo de límites n utilizado para trazar las zonas de severidad	76
Tabla 11: Criterios para la interpretación del ARL_{BSL}	78
Tabla 12: Gráficos de control p alternativos: Enfoque optimización de límites de control	83
Tabla 13: Síntesis de los resultados obtenidos para el GC p Corner-Fisher.....	86
Tabla 14: Síntesis de los resultados obtenidos para el GC p R&S.....	89
Tabla 15: Gráficos de control p alternativos: Enfoque uso de transformaciones	90
Tabla 16: Resultados ARL_{BSL} para GCs p basados en transformaciones.....	95
Tabla 17: Síntesis de los resultados para el GC p Kmod.....	100
Tabla 18. Síntesis de los resultados para el GC p Kmod y para los GCs p alternativos más destacados	101
Tabla 19: Gráficos de control u alternativos.....	110
Tabla 20: Síntesis de los resultados para el GC u Corner-Fisher.....	113
Tabla 21. Síntesis de los resultados para el GC u R&S.....	115
Tabla 22. Síntesis del rendimiento ARL para el GC u PT.....	118
Tabla 23: Síntesis de los resultados para el GC u Kmod.....	122

Tabla 24. Síntesis de los resultados para el GC u Kmod y para los GCs u alternativos más destacados 122

Tabla 25. Resultados del muestreo obtenidos en la segunda fase de muestreo 127

1.3 Índice de figuras

Figura 1. Distribución normal de la variación de X por causas aleatorias	9
Figura 2. Esquema de un gráfico de control para $X \sim N(\mu, \sigma)$	10
Figura 3. Ejemplo ilustrativo de un gráfico de control (GC)	11
Figura 4: Resumen esquemático de los gráficos de control uni-variantes más conocidos	12
Figura 5. Ejemplo de distribución simétrica para $B(100, 0.5)$	21
Figura 6. Ejemplo de distribución simétrica para $B(25,0.5)$	22
Figura 7. Ejemplo de distribución asimétrica positiva para $B(100,0.1)$	22
Figura 8. Ejemplo de distribución asimétrica positiva para $B(24,0.2)$	23
Figura 9. Ejemplo de distribución asimétrica positiva para $B(1000,0.05)$	23
Figura 10. Ejemplo de distribución asimétrica negativa para $B(100,0.9)$	24
Figura 11. Forma de la distribución normal o gaussiana	25
Figura 12. Histograma de los resultados del muestreo hecho para calcular p	30
Figura 13. Ejemplo de un GC np tipo Shewhart	33
Figura 14. Ejemplo de un GC p tipo Shewhart	34
Figura 15. Ejemplo de deterioro del proceso (incremento de p).....	35
Figura 16. Ejemplo de mejora del proceso (reducción de p)	35
Figura 17. GC p Shewhart - Comportamiento real de la colas alfa para $p = 0.1$	41
Figura 18. GC p Shewhart – Fluctuación del ARL_0 para $p = 0.1$	42
Figura 19. Ajuste de la aproximación normal a la binomial para $p= 0.1$ y $n=100$	43
Figura 20. Aproximación normal a la binomial: ajuste en las colas para $n = 100$ y $p = 0.1$	43
Figura 21. Probabilidad binomial acumulada para $n = 100$ y $p = 0.1$	44
Figura 22. Colas de la probabilidad binomial acumulada para $n = 100$ y $p = 0.1$	45
Figura 23. Curva ARL para GC p Shewhart con $n = 100$ y $p = 0.1$	47
Figura 24. Curvas ARL para $p = 0.05$ y $n = 220$ y 600	49

Figura 25. Probabilidad de Poisson para $c = 10$ y su aproximación normal.....	53
Figura 26. Colas de la probabilidad de Poisson para $c=10$ y su aproximación normal.....	54
Figura 27. Probabilidad de Poisson para $c = 20$ y su aproximación normal.....	54
Figura 28. Colas de probabilidad de Poisson para $c = 20$ y su aproximación normal	55
Figura 29. GC u Shewhart : Comportamiento real de la colas alfa para $u = 1$	60
Figura 30. GC u Shewhart: Oscilación del ARL_0	61
Figura 31. GC u Shewhart: Curvas ARL para $u = 1$ con $n = 11$ y $n = 18$	62
Figura 32. Relación entre la probabilidad de error Tipo I (α) y el ARL_0	66
Figura 33. GC p Shewhart: Variación del R_α en función de n para $p = 0.1$	68
Figura 34. GC p Shewhart: Variación del ARL-sesgo(%) para $p = 0.1$	68
Figura 35. GC p Shewhart: Relación entre ARL-sesgo(%) y R_α para $p = 0.1$	69
Figura 36: Bondad de ajuste de los valores ARL-sesgo(%) predichos por la ecuación (24) ..	72
Figura 37: Ejemplo de la relación entre ARL-sesgo(%) y ARL-ratio para $p = 0.05$	74
Figura 38. Gráfico de control p Shewhart: ARL-sesgo(%) para $p = 0.05$	75
Figura 39: Zonas de severidad del sesgo para curvas ARL de GCs p Shewhart	77
Figura 40. GC p Shewhart: Comportamiento del ARL_{BSL} para $p = 0.05$	80
Figura 41. GC p Shewhart: Relación entre R_α y ARL_{BSL} para $p = 0.05$	81
Figura 42. Ejemplo de colas alfa típicas para el GC p Corner-Fisher.....	84
Figura 43. Ejemplo del comportamiento típico del ARL_0 para el GC p Corner-Fisher	85
Figura 44. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Corner-Fisher	85
Figura 45. Ejemplo de colas alfa típicas para el GC p R&S.....	87
Figura 46. Ejemplo del comportamiento típico del ARL_0 para el GC p R&S.....	88
Figura 47. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p R&S	88
Figura 48. Ejemplo de colas alfa típicas para el GC p Arco Seno	91
Figura 49. Ejemplo del comportamiento típico del ARL_0 para el GC p Arco Seno	92
Figura 50. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Arco Seno.....	92

Figura 51. Ejemplo de curvas ARL con sesgo positivo y negativo	93
Figura 52. Ejemplo de colas alfa típicas para el GC p Kmod.....	98
Figura 53. Ejemplo del comportamiento típico del ARL_0 para el GC p Kmod.....	99
Figura 54. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Kmod	99
Figura 55. GC p Kmod: Relación entre el ARL_{BSL} y R_α para $p = 0.05$	102
Figura 56. GC p Kmod: Relación entre el ARL_{BSL} y R_α para $p = 0.10$	102
Figura 57. GC p Kmod: Ejemplo de curvas ARL para $p = 0.05$ con $n = 220$ y 600	104
Figura 58: Kmod p chart: Curvas ARL para $p = 0.05$ con $n = 220$ y 215	105
Figura 59. GC u Shewhart – Comportamiento del ARL_{BSL} para $u = 1$	107
Figura 60. GC u Shewhart – Comportamiento del ARL_{BSL} para $u = 2$	108
Figura 61. GC u Shewhart - Relación entre R_α y ARL_{BSL} para $u = 1$	109
Figura 62. Ejemplo de colas alfa típicas para el GC u Corner-Fisher.....	111
Figura 63. Ejemplo del comportamiento típico del ARL_0 para el GC u Corner-Fisher.....	112
Figura 64. Ejemplo de comportamiento típico del ARL_{BSL} para el GC u Corner-Fisher	112
Figura 65. Ejemplo de colas alfa típicas para el GC u R&S.....	114
Figura 66. Ejemplo del comportamiento típico del ARL_0 para el GC u R&S.....	114
Figura 67. Ejemplo del comportamiento típico del ARL_{BSL} para el GC u R&S	115
Figura 68. Ejemplo de colas alfa típicas para el GC u PT.....	116
Figura 69. Ejemplo del comportamiento típico del ARL_0 para el GC u PT.....	117
Figura 70. Ejemplo del comportamiento típico del ARL_{BSL} para el GC u PT.....	117
Figura 71. Ejemplo de colas alfa típicas para el GC u Kmod.....	120
Figura 72. Ejemplo del comportamiento típico del ARL_0 para el GC u Kmod.....	121
Figura 73. Ejemplo de comportamiento típico del ARL_{BSL} para el GC u Kmod	121
Figura 74. GC u Kmod: Relación entre el ARL_{BSL} y R_α para $u = 1$	124
Figura 75. GC u Kmod - Ejemplo de curvas ARL.....	125
Figura 76. Gráfico de control p Shewhart.....	127

Figura 77. Gráfico de control p K_{mod} 128

Figura 78. Curvas ARL para los gráficos p Shewhart y K_{mod} para $p = 0.1746$ y $n = 192$128

1.4 Índice de ecuaciones

(1)	9
(2)	9
(3)	29
(4)	32
(5)	36
(6)	36
(7)	36
(8)	36
(9)	37
(10)	37
(11)	37
(12)	37
(13)	39
(14)	51
(15)	51
(16)	52
(17)	53
(18)	57
(19)	57
(20)	57
(21)	58
(22)	58
(23)	60
(24)	71

(25)	72
(26)	77
(27)	83
(28)	87
(29)	90
(30)	91
(31)	93
(32)	94
(33)	97
(34)	110
(35)	113
(36)	116
(37)	119

1.5 Acrónimos

CEC	Control estadístico de la calidad
GC	Gráfico de control
GCs	Gráficos de control
LCI	Límite de control inferior
LCS	Límite de control superior
BC	Bajo control (relacionado a procesos)
FC	Fuera de control (relacionado a procesos)
ARL	Average Run Length – Longitud de tandas
ARL ₀	ARL cuando el proceso está bajo control
ARL ₁	ARL cuando el proceso está fuera de control
ARL _{BSL}	Valor cuantitativo de la severidad del sesgo de la curva ARL
UDI	Unidad de inspección (aplica a GC u)

1.6 Nomenclatura

p	Fracción de no-conformes de un proceso binomial
q	Fracción de conformes $q = 1 - p$
p_l	Fracción de no-conformes modificada
n	Tamaño de muestra
np	Número de no conformes
α_I	Cola alfa inferior – área de probabilidad por debajo del LCI
α_S	Cola alfa superior – área de probabilidad por encima del LCI
R_α	Ratio de las colas alfa (α_I/α_S)
c	Número promedio de defectos en una muestra
u	Promedio de defectos por unidad de inspección (UDI)
u_l	Promedio de UDIs modificado
$\lambda_{I S}$	Constante (I= inferior y S= superior) del factor Kmod utilizado para el nuevo gráfico de control p Kmod
$T_{I S}$	Constante (I= inferior y S= superior) del factor Kmod utilizado para el nuevo gráfico de control u Kmod

2 *Introducción y objetivos*

2.1 **Introducción**

Esta tesis resume el trabajo de investigación realizado para la obtención del título de Doctor en Estadística y Optimización por la Universidad Politécnica de Valencia. El trabajo fue llevado a cabo en el Departamento de Estadística de esa institución, dentro del equipo de investigaciones en “Técnicas estadísticas para la industria y la tecnología”.

Como se puede deducir del título de la tesis, el tema central de investigación fueron los gráficos de control por atributos, siendo éstos herramientas estadísticas que sirven para monitorizar procesos mediante atributos de calidad. La investigación en sí se enfocó en los gráficos del tipo uni-variante, específicamente en los basados en la distribución binomial (Gráficos p y np) y en los basados en la distribución de Poisson (Gráficos c y u).

Los gráficos por atributos Shewhart, que se fundamentan en la aproximación normal a las distribuciones binomial y de Poisson, están entre los más conocidos y forman parte imprescindible de cualquier libro de control de calidad. Además, también son de esencial enseñanza en todo curso de control estadístico de procesos, por lo que, basado en este antecedente, se puede decir con certeza que la mayoría de personas involucradas en el tema de calidad, tanto desde la academia como desde la industria, conocen acerca de estos gráficos y posiblemente muchos los hayan utilizado o recomendado.

Ahora bien, alguien que construyese un gráfico Shewhart siguiendo los pasos recomendados en la mayoría de libros de control estadístico, podría, con justa razón, asumir que ese gráfico sería idóneo para monitorizar el proceso para el cual fue diseñado, es decir, que el gráfico podría detectar de igual forma tanto mejoras como deterioros del proceso. Sin embargo, algo que pocas veces se menciona en la literatura, es que los gráficos por atributos Shewhart tienen una falencia que hace que su capacidad para detectar cambios en el proceso sea inadecuada. Esta falencia proviene de la errónea presunción de que la aproximación normal a las distribuciones binomial y de Poisson, es siempre apropiada para construir gráficos de control por atributos.

Sin querer entrar en aspectos complejos, cabe mencionar de momento, que la binomial y la de Poisson son distribuciones discretas altamente asimétricas y que debido a esto el ajuste de la aproximación normal es muy pobre en las colas de esas distribuciones, y que es ahí en donde

se origina la falencia antes mencionada. Fue justamente el estudio del impacto que tiene esta falencia en la capacidad de monitorización de los gráficos Shewhart, lo que dio inicio a la investigación.

Para poder llevar a cabo el estudio antes mencionado, se creó el parámetro ARL_{BSL} que sirve para cuantificar la capacidad de monitorización de gráficos por atributos. Los resultados obtenidos con el ARL_{BSL} para los gráficos Shewhart demostraron que estos tienen una capacidad de monitorización muy por debajo de lo deseado. Además, también se analizaron otros gráficos que habían sido propuestos como alternativas a los Shewhart, encontrándose que su capacidad de monitorización, aunque superior a los Shewhart, tampoco era satisfactoria.

Lo expuesto en el párrafo anterior motivó el desarrollo de dos nuevos gráficos de control, uno basado en la distribución binomial y otro en la de Poisson. Ambos gráficos son superiores a los ya existentes y además cuentan con un método muy simple para verificar su capacidad de monitorización.

Los resultados de la investigación, aparte de ser resumidos en esta tesis, han sido expuestos en congresos y también publicados en revistas científicas internacionales.

2.2 Planteamiento del problema

Dentro de las distintas fases que tomó la investigación está la considerada como la fase exploratoria, que abarca los análisis y estudios hechos para identificar el tema específico de investigación; los resultados de esta fase se resumen en la sección 4. En esa sección se describe la curva ARL, que sirve para evaluar la capacidad de monitorización de un gráfico de control, y también se presentan varias curvas ARL provenientes de gráficos de control por atributos del tipo Shewhart.

Ahora bien, un detalle que se observó en las curvas ARL de los gráficos Shewhart fue que todas eran sesgadas. Esto llamó mucho la atención ya que una curva ARL sesgada denota que la capacidad de monitorización del gráfico al cual pertenece es inadecuada, lo que hacía suponer que, si las curvas ARL de los gráficos Shewhart eran siempre sesgadas, entonces al contrario de lo que normalmente se asumía, estos gráficos de control no eran idóneos para monitorizar procesos. Este hecho nos llevó a la conclusión de que el estudio de la capacidad de monitorización de los gráficos de control Shewhart, era un tópico que merecía ser analizado con más profundidad.

Por otro lado, entre los varios parámetros analizados en la fase exploratoria estuvieron las “colas alfa”, que no son nada más que las probabilidades de que un punto caiga fuera de los límites del gráfico de control, cuando el proceso funciona de una manera estable. En la práctica lo habitual es simplemente asumir que estas colas, para los gráficos Shewhart, tienen siempre valores iguales a 0.00135, cabe mencionar que si este fuera el caso la curva ARL sería insesgada y que por ende el gráfico de control tendría una capacidad de monitorización idónea. Sin embargo, los análisis de las colas alfa demostraron que, al contrario de lo que normalmente se asume, sus valores no son iguales 0.00135 y que estos más bien varían y fluctúan (ver secciones 4.1.6.1 y 4.2.5.1).

En la fase exploratoria también se determinó que las variables binomial y de Poisson toman valores discretos específicos que están en función del tamaño de muestra (para p y c constantes respectivamente, ver secciones 4.1.7 y 4.2.3.1) y se concluyó que eso explicaba en buena parte el por qué las colas alfa no podían ser iguales o constantes. De esto se reflexionó que la naturaleza discreta de esas distribuciones debía afectar las colas alfa, no solo de los gráficos Shewhart, sino de todos los gráficos de control basados en esas distribuciones.

Tomando en cuenta el antecedente del sesgo en las curvas ARL y también el hecho de que las colas alfa no tienen un valor fijo y dado que estas influyen directamente en la severidad del sesgo de esas curvas, se razonó que el análisis de la capacidad de monitorización de gráficos por atributos por medio de curvas ARL era un tema que valía la pena investigar.

Una vez centrado el tema específico de investigación, se plantearon los objetivos que se detallan en la siguiente sección.

2.3 Objetivos de la investigación

2.3.1 Objetivo general

El objetivo general de la investigación es:

Desarrollar métodos para estimar y/o reducir el sesgo de la curva ARL en gráficos de control por atributos.

2.3.2 Objetivos específicos

Para lograr el objetivo general, se propusieron los siguientes objetivos específicos:

1. Desarrollar parámetros para cuantificar de forma efectiva el sesgo de las curvas ARL para gráficos de control por atributos.

2. Desarrollar métodos que faciliten la estimación de la capacidad de monitorización^a de gráficos de control por atributos.
3. Caracterizar la capacidad de monitorización de gráficos de control por atributos relevantes que hayan sido propuestos hasta la fecha de realización de la investigación.
4. Desarrollar nuevas propuestas de gráficos de control por atributos.
5. Publicar los resultados de la investigación en actas de congresos y en revistas científicas.

^a *En esta tesis, se entiende por capacidad de monitorización a la propiedad inherente y específica de cada gráfico de control, que define su sensibilidad para detectar cambios en el proceso.*

2.4 Estructura de la tesis

El presente documento se divide en ocho secciones que han sido realizadas siguiendo lineamientos generales de redacción de tesis, estas a su vez han sido sub-seccionadas cuidadosamente con el objetivo de resumir de la mejor forma el trabajo de investigación concerniente.

La sección 1 contiene aspectos que sirven al lector para navegar en los contenidos de la tesis, como lo son: el índice general, índice de figuras, índice de tablas etc. Por otro lado, en la sección 2 “Introducción y objetivos”, se da un preámbulo acerca del tema de investigación, se plantea el problema específico a investigar y se detallan los objetivos del trabajo.

En la sección 3 “Antecedentes respecto al tema de investigación”, se dan detalles generales acerca de gráficos de control uni-variantes, se describe su importancia, sus principios básicos, además de los grupos y tipos que existen. En esa sección también se identifican a los gráficos por atributos como el tema principal de esta tesis y se exponen los antecedentes que llevaron a plantear la pregunta que dio inicio a la investigación.

La sección 4 “Marco teórico base”, se divide en dos partes, la primera dedicada a gráficos basados en la distribución binomial y la segunda dedicada a gráficos basados en la distribución de Poisson, y contiene todos los estudios, tanto teóricos como prácticos, que llevaron a la identificación del tema específico en el que se enfocó la investigación, siendo este el análisis y caracterización de la capacidad de monitorización de gráficos por atributos a través de curvas ARL. En esta sección también se incluyen algunos métodos novedosos que sirvieron para identificar el tema específico antes mencionado.

La sección 5 “Resultados y contribuciones al conocimiento”, es la más importante de este documento ya que contiene los resultados de la investigación. La cantidad de contribuciones y resultados es significativa y en este punto no se los menciona para no redundar, solo cabe mencionar que al momento de redactar esta tesis la mayor parte de esos resultados ya han sido publicados en revistas científicas de prestigio internacional.

La sección 6 contiene las conclusiones generales acorde a los objetivos planteados, también se resumen las principales aportaciones obtenidas de la investigación y se proponen posibles temas para futura investigación.

La sección 7 contiene las referencias bibliográficas, mientras que la sección 8 “Anexos” contiene información adicional y complementaria.

3 Antecedentes respecto al tema de investigación

3.1 El rol de los gráficos de control en la obtención de la calidad

En la actualidad es una verdad indiscutible que empresas y organizaciones deben producir bienes y servicios de calidad para poder ser exitosas. Pero ¿Qué es la “calidad”? Y ¿Cómo se la obtiene?

Para dar una respuesta a estas preguntas vale la pena primeramente conocer cómo el concepto de “calidad” ha evolucionado en el tiempo. Un excelente resumen de esto lo dan Moreno-Luzón et al (2001), quienes muestran que las definiciones de calidad, desarrolladas principalmente a lo largo del siglo XX, podrían agruparse en las siguientes categorías:

- i. Calidad como conformidad a unas especificaciones
- ii. Calidad como satisfacción de las expectativas del cliente
- iii. Calidad como valor con relación al precio
- iv. Calidad como excelencia.

Shewhart (1939), Deming (1989), Ishikawa (1994) y Juran (1990), entre otros, han hecho grandes aportaciones hacia la evolución del concepto de la calidad. En este aspecto muy relevante resulta la aportación de David A Garvin, quien publica dos formidables artículos en 1984 y 1987. El primero lo titula *¿Qué significa verdaderamente la calidad de un producto?* (Garvin D. A., 1984), y el segundo *“Compitiendo en las Ocho Dimensiones Estratégicas de la Calidad”* (Garvin D. A., 1987). En estas publicaciones Garvin define a la calidad como el resultado de la sinergia de ocho factores (o dimensiones) y sienta las bases conceptuales para el desarrollo de lo que hoy en día se conoce como Gestión Estratégica de la Calidad Total (GECT). Las ocho dimensiones de Garvin se detallan en la Tabla 1.

Tabla 1: Las ocho dimensiones de la calidad de un producto según Garvin

Dimensiones		Lo que abarca cada dimensión
1	Rendimiento	Las características primarias y medibles del producto, sus especificaciones básicas.
2	Prestaciones	Aspectos adicionales al desempeño y que complementan el funcionamiento básico del producto.
3	Fiabilidad	La probabilidad de fallos, el tiempo de funcionamiento hasta que se produzca una avería.
4	Conformidad	El nivel de ajuste del producto a las especificaciones de diseño, es decir, el grado en que un producto, su proceso de elaboración y/o diseño se ajustan a estándares pre establecidos. Esta dimensión está relacionada al enfoque de calidad conocido como cumplimiento de especificaciones.

5	Durabilidad	Es el periodo de vida útil del producto.
6	Capacidad y calidad de servicio	Atención al cliente, rapidez y eficacia para resolver quejas y reclamos, eficiencia y costos de reparación.
7	Estética	Es la apreciación subjetiva del producto hecha través de los sentidos. Requiere conocer y comprender aspectos como el entorno cultural y la edad de los potenciales clientes o consumidores.
8	Calidad percibida	Es otra dimensión subjetiva, que toma en cuenta el juicio que hace un consumidor o cliente respecto a un producto basándose en aspectos como publicidad, comentarios, experiencia propia, prestigio de la marca etc.

Como se puede apreciar de la Tabla 1, llegar a obtener “calidad” no es tarea fácil y según Moreno-Luzon et al (2001), para lograrlo se requiere de la apropiada implementación de herramientas y técnicas para su gestión y control. Es en este último punto donde el Control Estadístico de la Calidad (CEC) desempeña un papel primordial.

El CEC comprende un conjunto de herramientas estadísticas que han sido desarrolladas con la finalidad de facilitar la obtención de la calidad, principalmente a través del control y la mejora de procesos. Ha sido demostrado también que el CEC puede ayudar a aumentar la rentabilidad de una empresa (Garvin D. A., 1987, pág. 37), ya que mejorando la dimensión “Conformidad”, detallada en la Tabla 1, se pueden reducir los costes de manufactura y reparación de un producto.

Dentro de lo que se conoce como las siete herramientas clásicas de la calidad, ver Miranda-González et al (2007, pág. 75), se incluyen varios métodos estadísticos entre los que destacan: i) La hoja de registro ii) El histograma iii) El diagrama de Pareto iv) El diagrama de dispersión y v) El gráfico de control. Estas herramientas son de esencial conocimiento por lo que es de esperarse que la mayoría de personas involucradas en Gestión de la Calidad las conozcan y utilicen.

Sin embargo, el CEC va mucho más allá de las herramientas antes mencionadas. El libro de A.J. Duncan (1986) titulado “Control de Calidad y Estadística Industrial”, y que ha sido publicado desde los años 50’s del siglo XX, muestra la diversa gama de métodos y técnicas estadísticas que ya abarcaba el CEC a mediados el siglo pasado, siendo las principales: 1) Muestreo para aceptación de lotes 2) Inspección rectificadora 3) Gráficos de control 4) Pruebas de hipótesis 5) Análisis de la varianza y covarianza 6) Regresión y correlación y 7) Diseño de experimentos. Es interesante notar que todas estas técnicas todavía son relevantes en la actualidad.

Aparte de Duncan, y a lo largo de los años, otros autores han publicado obras dedicadas enteramente al CEC, siendo algunos de ellos: Montgomery (2009), Ryan (2011), Mitra (2008) y Qiu (2014). Cabe destacar que el CEC está en constante avance, siempre tratando de mejorar los métodos y técnicas existentes, así como en busca de soluciones a los retos impuestos por las nuevas tecnologías de producción, ejemplo de esto es el advenimiento de una rama de CEC dedicada exclusivamente a procesos multi y mega variantes. Existen varios libros dedicados a estos últimos temas, de entre los que destacan los publicados por Fuchs&Kenett (1998) y Eriksson et al (2013).

En los párrafos anteriores se ha destacado, de una manera breve y puntual, el importante rol que tiene el CEC en la obtención de la calidad y se ha mencionado brevemente los diversos métodos estadísticos que abarca. Sin embargo, de entre todos aquellos métodos el *Gráfico de Control* destaca como uno de los más importantes debido a que ha sido utilizado desde los inicios del CEC y porque su utilidad hacia la obtención de la calidad está más que comprobada (Shewhart, 1939). Es justamente debido a su transcendencia, la razón por la cual se lo escogió como tema central de la investigación que se resume en este documento.

En la siguiente sección se da una reseña acerca de gráficos de control y se introduce el tema específico de esta tesis.

3.2 Gráficos de control

3.2.1 Generalidades

Un gráfico de control (GC¹) es un método estadístico utilizado principalmente para monitorizar procesos de producción de bienes y servicios, siendo su principal función la de identificar si estos están o no bajo control.

Se considera que un proceso está bajo control (BC) si la variabilidad del producto producido es provocada por causas aleatorias inherentes al proceso, mientras que como fuera de control (FC) cuando la variabilidad se debe principalmente a causas atribuibles que no son inherentes al proceso, ejemplo de estas podrían ser: diferencias entre máquinas, entre trabajadores y entre materiales.

Para profundizar un poco en los principios estadísticos del GC, vamos a suponer que se toman muestras de productos de un proceso en intervalos más o menos regulares y que el tamaño de

¹ GC(s) = Gráfico(s) de Control

la muestra (n) es adecuado. Ahora bien, de cada muestra se podría calcular el valor de un estadístico X , el cual podría ser: el promedio, la amplitud, la fracción de no-conformes etc. En este caso X_i ($i = 1, 2, 3\dots$) sería una variable aleatoria sometida a las fluctuaciones del muestreo.

Ahora bien, sí el proceso está bajo control estadístico, la distribución de X_i debería tomar una forma estadística determinada. Suponiendo que la distribución de X_i es normal, como se muestra en la Figura 1, entonces el promedio (\bar{X}_i) y la varianza (S_i^2) muestral darían una estimación aproximada de la media (μ) y de la desviación estándar (σ) reales de X . Con esa información se podrían calcular las siguientes probabilidades:

- i. La probabilidad (P_1) de que X_i tome un valor que este dentro de los límites LCL y UCL², ver Figura 1. En este caso P_1 es calculado mediante la ecuación (1).
- ii. La probabilidad (P_2) de que X_i tome un valor que exceda cualquiera de los límites LCL o UCL. En este caso P_2 es calculado mediante la ecuación (2).

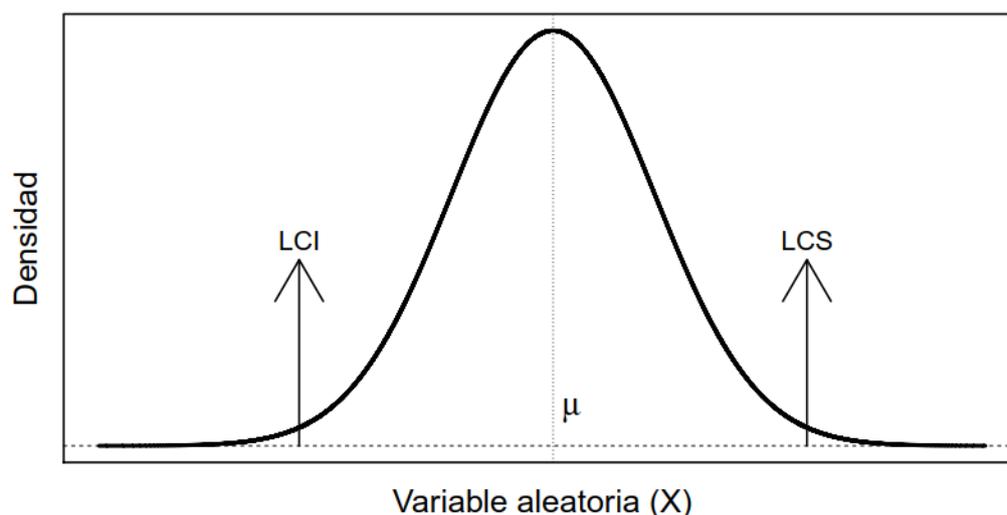


Figura 1. Distribución normal de la variación de X por causas aleatorias

$$P_1(LCI \leq X_i \leq LCS) = \int_{LCI}^{LCS} f(x) dx \quad (1)$$

$$P_2(LCI > X_i > LCS) = 1 - \int_{LCI}^{LCS} f(x) dx \quad (2)$$

Dónde:

² LCS|LCI= Limite de Control Superior|Inferior

$$\text{Función de densidad para una distribución normal} = f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Una vez establecidos los valores aproximados de μ y σ , juntos con los límites y las probabilidades (i) y (ii), es posible construir un GC en donde se trazarían los valores de X_i y cuyos límites (LCL y UCL) servirían de referencia para valorar si el proceso está o no bajo control. Cabe mencionar que es una práctica universalmente aceptada el colocar los límites de control a una distancia de ± 3 desviaciones estándar ($\pm 3\sigma$) de la media (μ).

La Figura 2 muestra las características fundamentales que tendría ese gráfico, como se puede observar el eje vertical corresponde a los valores que podría tomar X_i , mientras que el eje horizontal corresponde al número o identificación de cada muestra, y que van ordenados en una secuencia acorde a la evolución de la producción en el tiempo. La Figura 2 incluye la distribución de X como una ayuda visual para apreciar la relación que tiene con los límites de control y la línea promedio.

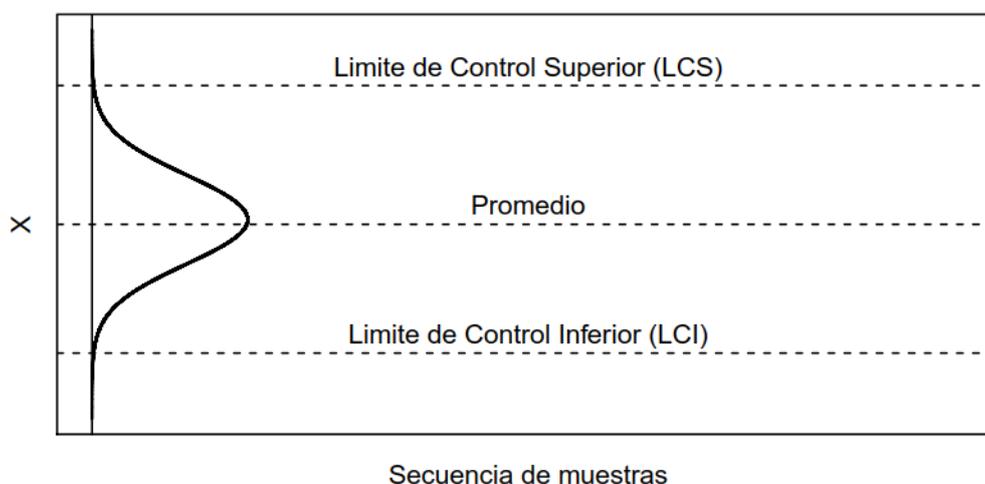


Figura 2. Esquema de un gráfico de control para $X \sim N(\mu, \sigma)$

La Figura 3 muestra un ejemplo de un gráfico de control donde X es la fracción de no conformes, este tipo de gráfico se analiza a profundidad en secciones posteriores y por ahora solo sirve para ilustrar los fundamentos teóricos del GC. Como se puede apreciar ningún valor de X_i rebasa los límites de control, por lo que se podría asumir que el proceso está bajo control estadístico. Sin embargo, aparte de verificar si existen puntos por fuera de los límites, también es importante analizar la distribución de tandas (también denominadas rachas o series) de X_i en el gráfico.

Según Montgomery (2009, pág. 195), para asumir que el proceso está bajo control las tandas de X_i trazadas en el gráfico de control (GC), aparte de caer dentro de sus límites, deben también tener una distribución aleatoria. Distribuciones de X_i que podrían sugerir un comportamiento no aleatorio del proceso podrían ser una tanda por encima o por debajo de la media, así como una tanda con valores ascendentes o descendentes.

Volviendo a la Figura 3, a primera vista y ya que todos los puntos están dentro de los límites, se podría asumir que el proceso está bajo control. Sin embargo, la tanda que abarca las muestras de la 13 a la 22, denota un patrón no aleatorio donde todos los valores de X_i están por debajo de la media; esto podría ser indicativo de que alguna causa asignable está afectando el proceso y que por lo tanto éste se encuentre fuera de control.

La teoría de distribución de tandas, o en inglés “run rules”, es bastante amplia y no forma parte de la investigación resumida en este documento, razón por la cual no se profundiza en el tema. No obstante, varios autores han publicado varios trabajos interesantes al respecto, entre otros: Duncan (2000, pág. 418), Nelson (1984), Champ&Woodhall (1987) y Khoo (2003).

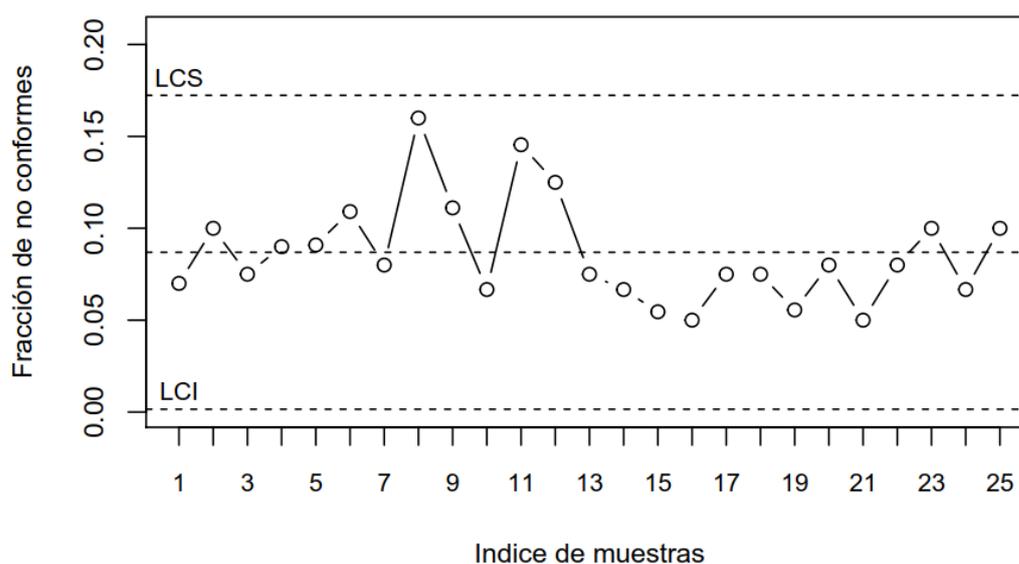


Figura 3. Ejemplo ilustrativo de un gráfico de control (GC)

Como se ha mostrado en los párrafos anteriores, mediante la aplicación de principios estadísticos simples es posible crear gráficos de control que pueden servir para monitorizar procesos de producción. Sin embargo, la gran diversidad de bienes y servicios producidos en la industria moderna requieren a su vez de una extensa diversidad de tipos de procesos para producirlos, lo que ha dado lugar a que se desarrollen una gran variedad de gráficos de control. En la siguiente sección se detallan algunos de ellos y se identifica el tipo específico del que trata esta tesis.

3.2.2 Tipos de gráficos de control

En general los Gráficos de Control (GCs) se dividen en dos grupos de acuerdo a la naturaleza de los datos obtenidos del proceso a controlar y que determinan el estadístico utilizado para monitorizarlo. Estos grupos son: i) Gráficos de control por variables y ii) Gráficos de control por atributos. Estos a su vez podrían separarse en uni-variantes, cuando se utiliza solo una variable o atributo, y multi-variantes cuando se utilizan varias variables o atributos.

Los gráficos de control multi-variantes están fuera del alcance de la investigación hecha para esta tesis, por lo que no se ahonda en ese tema. Sin embargo, varios autores han publicado trabajos que resumen los diversos tipos de gráficos multi-variantes desarrollados en estas últimas décadas, por nombrar algunos, Lowry & Montgomery (1995), Kourti & MacGregor (1996), Bermisi et al (2006), Yeh et al (2006), Topalidou & Psakaris (2009), y Dhini & Surjandari (2016).

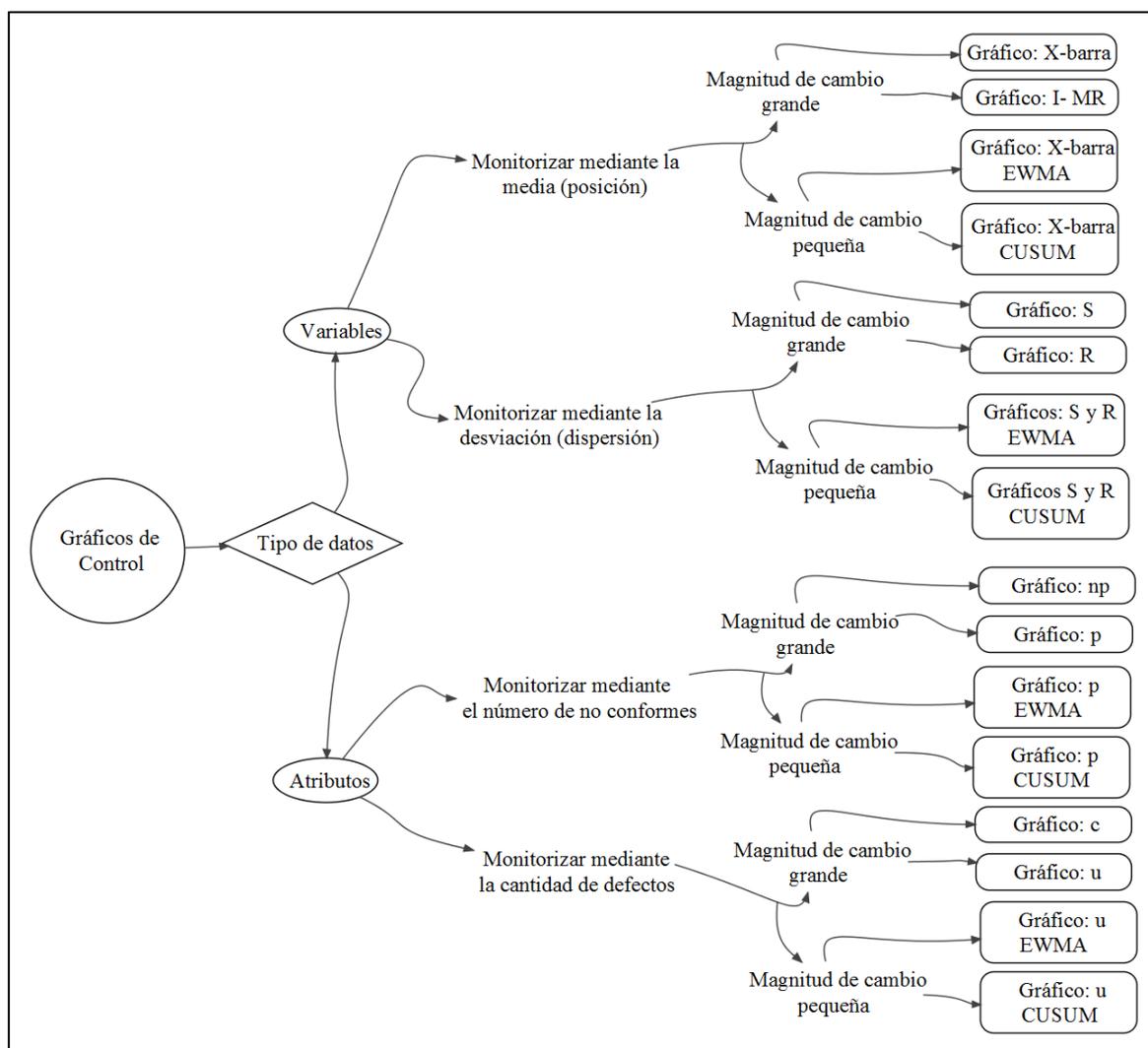


Figura 4: Resumen esquemático de los gráficos de control uni-variantes más conocidos

La Figura 4 muestra un esquema donde constan los gráficos uni-variantes más conocidos, y que han sido clasificados de acuerdo al tipo de datos que se utilizan para obtener el estadístico (\bar{X}) y a la magnitud de cambio que se desea detectar. Cabe señalar que cada gráfico requiere de la aplicación de técnicas estadísticas específicas, sin embargo hay características fundamentales, como el uso de límites de control, que son comunes para todos.

En la siguiente sección se realiza una breve reseña de los GCs uni-variantes más conocidos y se identifican los que incumben a esta tesis.

3.2.2.1 Gráficos de control uni-variantes basados en variables continuas

La principal peculiaridad de estos gráficos, es que sus estadísticos están basados en características de calidad que pueden ser valoradas a través de mediciones hechas en escalas numéricas continuas. Ejemplos de este tipo de variables son el peso, la longitud, el voltaje, la resistencia mecánica etc. En este caso, son dos los estadísticos que podrían ser usados para controlar el proceso, siendo estos: los estadísticos de posición y de dispersión.

Cuando se monitoriza la posición mediante el promedio (o media), se utiliza el gráfico \bar{X} (X-barra), mientras que cuando se monitoriza la variabilidad, dependiendo del caso, se podrían utilizar el gráfico R, que denota las variaciones en los rangos de las muestras, o el gráfico S, que denota las variaciones de las desviaciones estándar de las muestras. Por otro lado, cuando el proceso genera series de elementos individuales que no pueden ser agrupados en subgrupos, se utiliza el gráfico de control para elementos individuales I-MR. Montgomery (2009, pág. 226) analiza a profundidad todos estos gráficos.

Los gráficos \bar{X} , R y S sirven principalmente para detectar magnitudes de cambios en el proceso que van de moderadas a grandes. Para detectar cambios pequeños existen los gráficos EWMA (Exponentially Weighted Moving Average) y CUSUM (Cumulative Sum), que fueron propuestos originalmente por Roberts (1959) y Page (1954) respectivamente. De estos gráficos existen variaciones interesantes, como los planteados por Chen et al (2004) y Zhang&Yu (2005), que monitorizan la media y la variabilidad en un solo gráfico.

3.2.2.2 Gráficos de control uni-variantes basados en atributos de calidad

Es muy común encontrar procesos de producción en los que, para determinar la conformidad del producto, en vez de variables continuas se utilizan variables discretas. En general las variables discretas están relacionadas con los “*atributos de calidad*” del producto. Ejemplo de atributos son:

- 1) Aquellas características de calidad que no pueden ser evaluadas mediante escalas numéricas continuas, como por ejemplo: tono de color, raspaduras, partes faltantes, errores en relleno de formularios, número de errores médicos en un hospital, número de quejas por semana etc.
- 2) Aquellas características de calidad que sí podrían ser evaluadas mediante escalas numéricas continuas pero que debido al tiempo, costo o necesidad, son evaluadas mediante técnicas no exactas. Por ejemplo, las dimensiones de un artículo podrían ser parámetros importantes para determinar su calidad, sin embargo, en vez de medirlas se podría hacer uso de galgas para verificar si estas están dentro de un rango o estándar preestablecido.

Los GCs por atributos fundamentan sus estadísticos en las características de calidad antes mencionadas. Los cuatro GCs uni-variantes por atributos más conocidos son: Los gráficos np y p , los cuales están basados en el número de no conformes por muestra y los gráficos c y u que se basan en el número de defectos por unidad de inspección. Cabe señalar que esta tesis se enfoca en los **gráficos por atributos uni-variantes**, razón por la cual de este punto en adelante, cuando se mencione GCs por atributos, se referirá específicamente a los del tipo uni-variante.

Los GCs por atributos se aplican tanto en la industria de bienes como en la de servicios, aunque es en esta última donde tienen una relevancia especial, ya que allí existen procesos que solo pueden ser monitorizados mediante este tipo de gráficos. Ejemplos de su uso y aplicación pueden ser vistos en: Montgomery (2009, pág. 288), Mitra (2008, pág. 265), (Carey, 2003), Morton et al (2002), Wood (1994), Amin (2001) y Chen et al (2015). Para hacerlos más sensibles a cambios de magnitud pequeños, también existen variaciones EWMA y CUSUM, ejemplos de estos últimos pueden verse en Lucas (1985), Gan (1993) y Trevanich & Bourke (1993).

Debido a su importancia, el conocer acerca de los GCs por atributos es esencial para todo aquel que desee trabajar en control de calidad. Los GCs por atributos más populares y que son de imprescindible estudio y conocimiento son los del tipo Shewhart, cuya característica principal es que sus límites de control se determinan mediante ecuaciones basadas en la aproximación normal a las distribuciones binomial y de Poisson.

Ahora bien, es conocido que la aproximación normal a las distribuciones antes mencionadas no siempre da resultados satisfactorios. Ejemplo de esto lo dan Agresti & Coull (1998) quienes demuestran, desde el punto de vista de construcción de intervalos de confianza, que la

aproximación normal a la binomial generalmente da resultados muy pobres. Otros trabajos que confirman esos resultados son los publicados por Agresti & Brian (2000), Brown et al (2001) y Brown et al (2002). Por otro lado, también hay trabajos, como el de Rich (2009), que demuestran que la aproximación normal a la distribución de Poisson no es del todo satisfactoria.

Es muy común encontrar en los textos de estadística, reglas empíricas con las que supuestamente se obtienen las condiciones para que la aproximación normal a las distribuciones antes mencionadas, sea satisfactoria. Probablemente la mayoría de los que hemos estudiado control de calidad hemos visto estas reglas, ejemplos de ellas son i) $np \geq 5$ (o 10) para la distribución binomial (Ryan T. P., 2011, pág. 182) y ii) $C \geq 10$, para la distribución de Poisson (Mitra, 2008, pág. 212). El significado de los parámetros np y C se los detalla en las secciones 4.1 y 4.2 respectivamente.

En general, la existencia de estas reglas puede llevarnos a asumir que, siempre y cuando los GCs por atributos tipo Shewhart se adhieran a ellas, estos serán apropiados para monitorizar procesos. Sin embargo, existen trabajos como los de Schader & Schmid (1989) y Chang et al (2008) en donde se demuestra que estas reglas no garantizan la obtención de una adecuada aproximación a la normal; por lo que es de sentido común asumir que tampoco aseguren la obtención de GCs idóneos.

De lo expuesto en los párrafos anteriores nació la pregunta que dio inicio a la investigación que se resume en esta tesis. La pregunta fue: *¿Cómo afecta la imprecisión de la aproximación normal a la capacidad de monitorización de los CGs por atributos tipo Shewhart?*

La búsqueda de una respuesta a esta pregunta dio lugar a que surgieran otras interrogantes, que a su vez, dieron origen a los objetivos propuestos para la investigación.

4 Marco teórico base

4.1 Gráficos de control por atributos basados en la distribución binomial

Al final de la sección 3.2.2.2 se mencionó que la pregunta que dio lugar al trabajo de investigación resumido en esta tesis fue: *¿Cómo afecta la imprecisión de la aproximación normal a la capacidad de monitorización de los GCs por atributos tipo Shewhart?* En esta sección se presentan los fundamentos teóricos que sirvieron de base para dar respuesta a esa pregunta, para el caso específico de GCs basados en la distribución binomial.

Cabe mencionar que lo expuesto en esta sección, incluye resultados de lo que se considera como la fase exploratoria del tema de investigación. Al comienzo de esa fase, y como es normal en un proceso de investigación, todavía no se tenía claro si el trabajo de exploración impulsado por la pregunta expuesta en el párrafo anterior, daría lugar a la identificación de un tema novedoso digno de una tesis doctoral. Afortunadamente, al final de esa fase se estableció que el tema del sesgo en las curvas ARL tenía potencial, por lo que se decidió enfocar el trabajo en ese tema y fue en ese momento cuando se definieron los objetivos que se exponen en la sección 2.3.

4.1.1 Introducción al uso de la distribución binomial en gráficos de control por atributos

Vamos a suponer que un cierto tipo de artículo es producido por medio de un proceso estable y que de este se toman muestras independientes, de tamaño n , y que cada uno de los artículos incluidos en las muestras es inspeccionado y clasificado como “conforme” o “no-conforme” de acuerdo a atributos de calidad previamente establecidos. Vamos a suponer también que la fracción promedio de no-conformes por muestra, p , ha sido establecida.

En este caso el número de artículos no-conformes por muestra, denotado como x_i ($i = 1, 2, 3, \dots$) podría ser modelado como una variable aleatoria que sigue una distribución de probabilidad binomial $B(n, p)$. Por lo que el proceso en principio podría ser monitorizado a través de un gráfico de control estadístico basado en esa distribución.

4.1.2 Propiedades de la distribución binomial $B(n, p)$

La función de probabilidad $P(x)$ de la distribución binomial está dada por la ecuación (a.1). Desde el punto de vista de control de procesos, mediante esta función se obtiene la probabilidad de hayan x artículos no conformes en una muestra con n observaciones. La probabilidad de que x artículos sean clasificados como no conformes está dado por p , y como conformes por $q = 1 - p$. La clasificación se la realiza en base a atributos de calidad preestablecidos y se asume que el muestreo es realizado de forma aleatoria e independiente.

$$P(x) = Q(n, x)p^x q^{n-x}, \quad (a.1)$$

$$0 \leq p \leq 1,$$

$$x = 0, 1, 2 \dots n$$

$$\text{Dónde: } Q(n, x) = \frac{n!}{x!(n-x)!}$$

$$q = 1 - p$$

Ahora bien, estimar la media $E(x)$ y la varianza $Var(x)$ de la distribución binomial mediante ecuaciones simples es de fundamental importancia. A continuación se desarrollan expresiones matemáticas para $E(x)$ y $Var(x)$ mediante los siguientes métodos: i) Expansión binomial y ii) Generación de momentos.

4.1.2.1 Media y Varianza de la distribución binomial mediante expansión binomial

Sé conoce que $(p + q)^n$ puede ser representada como una expansión binomial, tal como se muestra en la ecuación (a.2):

$$(p + q)^n = \sum_{x=0}^n Q(n, x)p^x q^{n-x}, \quad 0 \leq p \leq 1 \quad (a.2)$$

Tomando la primera derivada de (a.2) con respecto a p y luego multiplicando esa expresión por p nuevamente, se obtiene (a.3):

$$p \frac{d}{dp} (p + q)^n = \sum_{x=0}^n x Q(n, x) p^x q^{n-x} \quad (a.3)$$

Resolviendo la parte derecha de (a.3) se obtiene (a.4):

$$np(p + q)^{n-1} = \sum_{x=0}^n x Q(n, x) p^x q^{n-x} \quad (a.4)$$

Si en (a.4) dejamos que $q = 1 - p$, se tiene como resultado (a.5):

$$np = \sum_{x=0}^n x P(x) \quad (\text{a.5})$$

En (a.5) es posible comprobar que $\sum_{x=0}^n x P(x)$ es el valor aproximado del promedio, $E(X)$, de la distribución binomial $B(n, p)$, por lo que $E(X)$ también puede ser estimado mediante (a.6).

$$E(x) = \bar{x} = np \quad (\text{a.6})$$

Para encontrar una expresión para calcular la varianza de la distribución binomial, se podría tomar la segunda derivada de (a.2) con respecto a p y luego multiplicar esa expresión por p^2 , esto da (a.7):

$$p^2 \frac{d^2}{dp^2} (p + q)^n = \sum_{x=0}^n x(x-1)Q(n, x)p^x q^{n-x} \quad (\text{a.7})$$

Resolviendo la parte izquierda de (a.7) se obtiene (a.8):

$$n(n-1)p^2(p+q)^{n-2} = \sum_{x=0}^n x(x-1)Q(n, x)p^x q^{n-x} \quad (\text{a.8})$$

Si en (a.8) dejamos que $q = 1 - p$, se tiene como resultado (a.9) y consecuentemente (a.10) que es el promedio $E(x)^2$:

$$n(n-1)p^2 = \sum_{x=0}^n x^2 P(x) - \sum_{x=0}^n x P(x) \quad (\text{a.9})$$

$$n(n-1)p^2 + \sum_{x=0}^n x P(x) = \sum_{x=0}^n x^2 P(x)$$

$$n(n-1)p^2 + np = \sum_{x=0}^n x^2 p(x) = E(x)^2 \quad (\text{a.10})$$

La expresión de la varianza se la obtiene mediante usando (a.10) y (a.6):

$$\begin{aligned} \text{Var}(x) &= E(x)^2 - [E(x)]^2 \\ \text{Var}(x) &= (\text{a.10}) - (\text{a.6})^2 \\ \text{Var}(x) &= n(n-1)p^2 + np - (np)^2 \\ \sigma^2 &= \text{Var}(x) = np(1-p) \end{aligned} \quad (\text{a.11})$$

4.1.2.2 Media y Varianza de $B(n, p)$ mediante generación de momentos

Si X es una variable aleatoria (v.a), entonces el momento de orden k de X es definido como:

$$E(X^k)$$

La función generadora de momentos de una v.a. X discreta, es una función de valores reales $M_x(t)$ definida como se muestra en la ecuación (b.1):

$$M_x(t) = \sum_{x=0}^n e^{xt} P(x) \quad (\text{b.1})$$

A (b.1) se la denomina como función generadora de momentos, porque estos pueden ser obtenidos derivando esta función k veces y evaluando esas derivadas en $t=0$, tal como lo establece en (b.2).

$$E(X^k) = \frac{\partial}{\partial t^k} M_x(t)|_{t=0} \quad (\text{b.2})$$

Dónde:

- El primer momento de $E(X^1)$ da la media de la distribución
- El segundo momento $E(X^2)$ puede ser usado para encontrar la varianza de la distribución mediante $V(x) = E(X^2) - (E(X^1))^2$

Ahora bien, para obtener $M_x(t)$ para la distribución binomial reemplazamos (a.1) en (b.1) para obtener (b.3):

$$\begin{aligned} M_x(t) &= \sum_{x=0}^n e^{xt} Q(n, x) p^x q^{n-x} \\ &= \sum_{x=0}^n (pe^t)^x Q(n, x) q^{n-x} \end{aligned} \quad (\text{b.3})$$

Finamente (b.3) puede ser representada como la expresión binomial (b.4)

$$M_x(t) = (q + pe^t)^n \quad (\text{b.4})$$

Reemplazando (b.4) en (b.2) y resolviendo las derivadas con $t=0$, se obtienen fácilmente los parámetros característicos de la distribución binomial, como la media (b.5) y la varianza (b.6).

$$E(X^1) = np(q + p)^{n-1} = np = \bar{x} = \text{posición o media de la distribución} \quad (\text{b.5})$$

$$E(X^2) = n(n-1)p^2 + np$$

Dado que la Varianza es igual a: $V(x) = E(X^2) - (E(X))^2$, entonces:

$$V(x) = [n(n-1)p^2 + np] - (np)^2$$

$$V(x) = np(1-p) \quad (b.6)$$

4.1.2.3 Simetría de la distribución binomial

La simetría sirve para describir la manera en que la variable aleatoria (X) se distribuye con respecto a un valor central. En general para una distribución discreta como la binomial, la simetría se la precisa con relación al valor promedio. Destacan los siguientes tipos de simetrías:

- i. **Distribución totalmente simétrica:** Se da cuando las frecuencias de distribución de la variable aleatoria (X), son iguales a ambos lados de la media (ver Figura 5).
- ii. **Distribución asimétrica positiva:** Se da cuando las frecuencias más altas se ubican a la izquierda de la media y las más bajas a la derecha, presentando un alargamiento en la cola derecha de la distribución (ver Figura 7)
- iii. **Distribución asimétrica negativa:** Se da cuando las frecuencias más altas se ubican a la derecha de la media y las más bajas a la izquierda, presentando un alargamiento en la cola izquierda de la distribución (ver Figura 10).

Existen varios índices que sirven para describir o determinar las características de asimetría de una distribución. Un índice bastante conocido es el de Fisher, que es descrito por la ecuación (c.3):

$$\gamma = \left[\frac{E(X - u)^3}{\sigma^3} \right] \quad (c.3)$$

Dónde:

- $\mu = \text{media}$
- $\sigma = \text{Desviación estándar}$

La versión de ecuación (c.3) con base en la función generadora de momentos de la distribución binomial es igual a (c.4).

$$\gamma = \frac{\mu_3}{\mu_2^{3/2}} \quad (c.4)$$

Donde μ_n es el $n.th$ momento central: $\mu_n = E(X - \bar{x})^n$

Resolviendo (c.4) se obtiene (c.5), que es el índice de asimetría de Fisher para la distribución binomial.

$$\gamma = \frac{1 - 2p}{\sqrt{np(1-p)}} \quad (\text{c.5})$$

Dónde:

- $\gamma = 0$, indica una distribución totalmente simétrica
- $\gamma > 0$, indica una distribución asimétrica positiva
- $\gamma < 0$, indica una distribución asimétrica negativa

De la Figura 5 a la Figura 10 se muestran ejemplos de las distintas asimetrías antes mencionadas. Estas sirven para demostrar las siguientes características de simetría que son inherentes para la distribución binomial:

- i. Cuando $p = 0.5$, la distribución binomial es simétrica para cualquier tamaño de muestra.
- ii. Cuando $p \ll 0.5$, la distribución binomial tiene asimetría positiva, aún para tamaños de muestra (n) grandes.
- iii. Cuando $p \gg 0.5$, la distribución binomial tiene asimetría negativa.

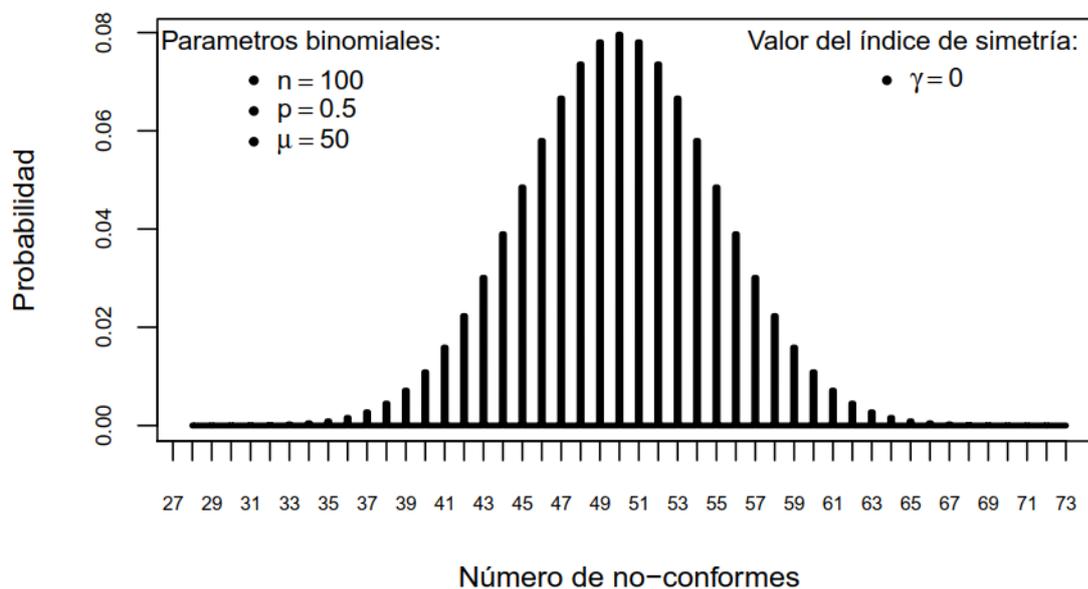
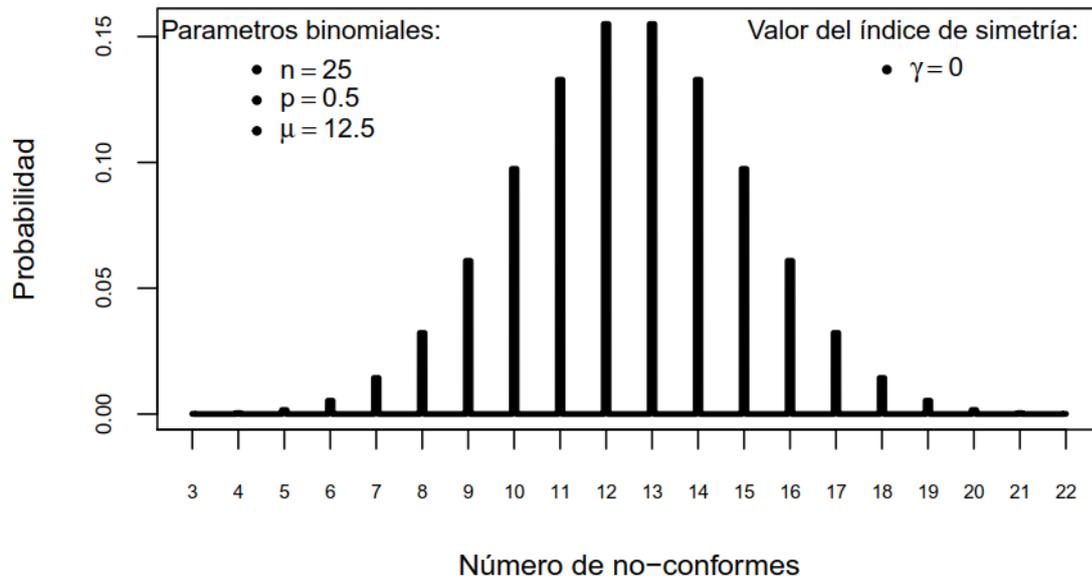
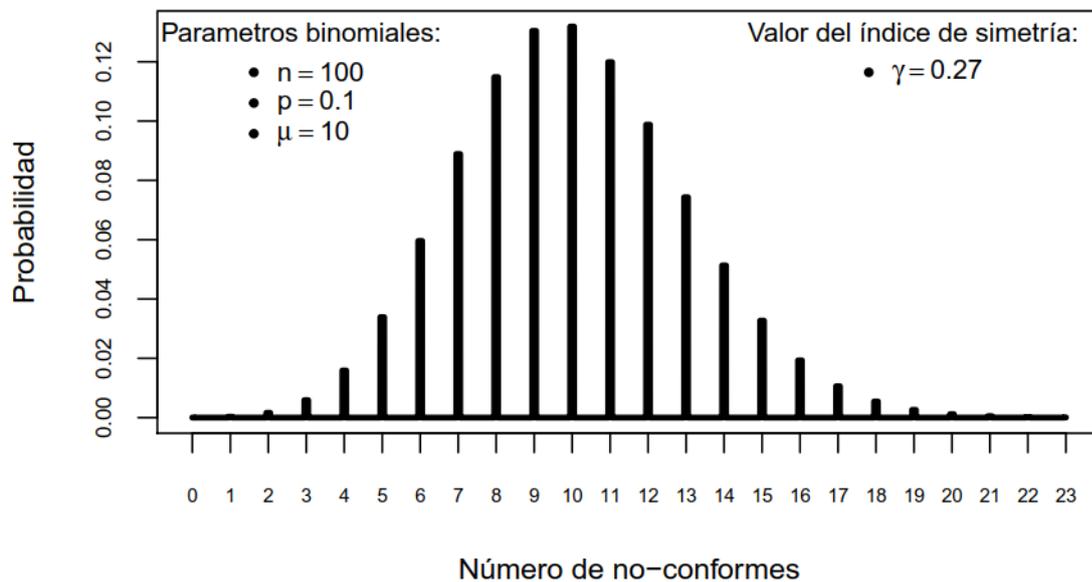
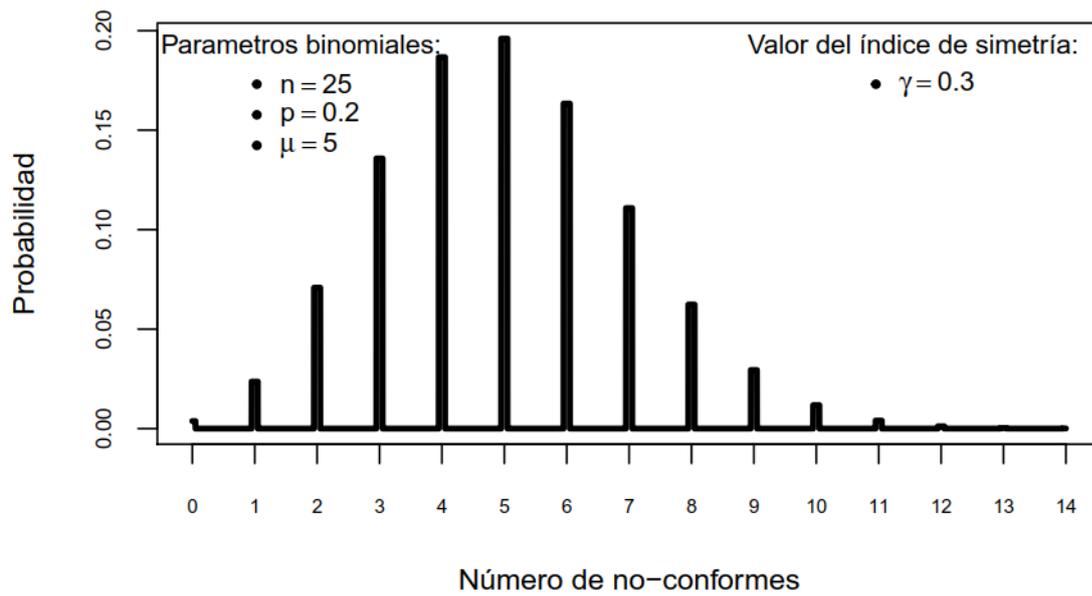
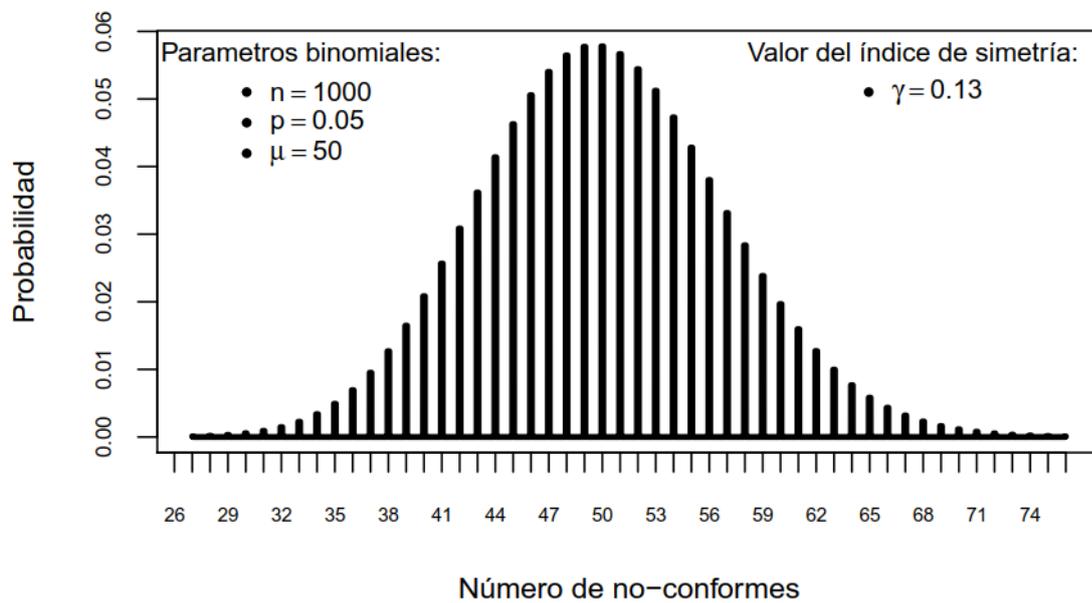


Figura 5. Ejemplo de distribución simétrica para $B(100, 0.5)$

Figura 6. Ejemplo de distribución simétrica para $B(25,0.5)$ Figura 7. Ejemplo de distribución asimétrica positiva para $B(100,0.1)$

Figura 8. Ejemplo de distribución asimétrica positiva para $B(24,0.2)$ Figura 9. Ejemplo de distribución asimétrica positiva para $B(1000,0.05)$

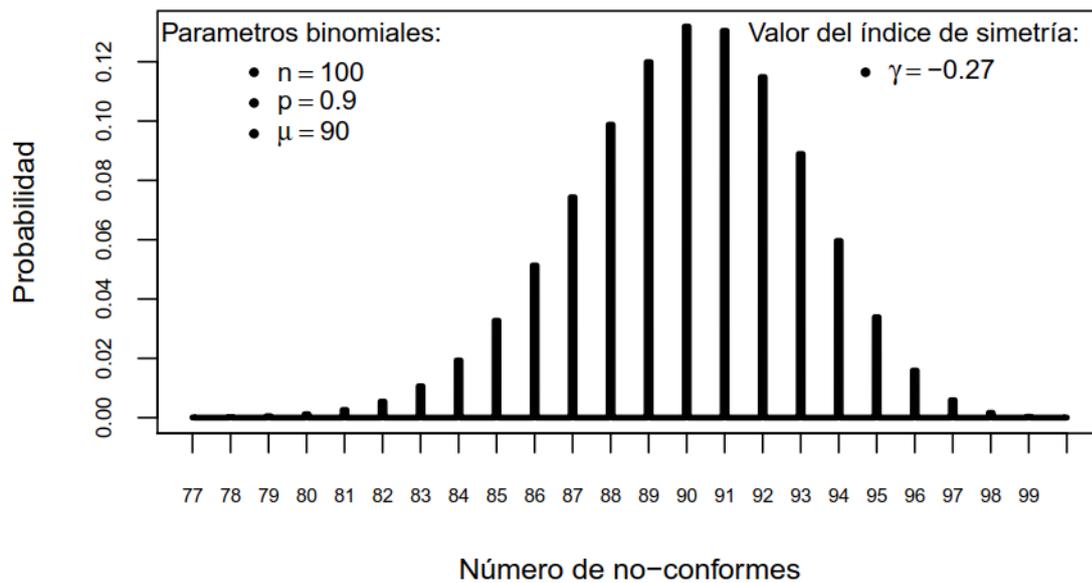


Figura 10. Ejemplo de distribución asimétrica negativa para $B(100,0.9)$

4.1.2.4 Sinopsis de la sección 4.1.2.

En esta sección se han desarrollado expresiones matemáticas para la media, varianza y sesgo de la distribución binomial. Por lo que se puede afirmar lo siguiente de una variable aleatoria que sigue una distribución binomial:

$$\text{Cuando: } X \sim B(n, p)$$

Entonces:

$$\bar{x} \equiv np$$

$$\text{Var}(x) \equiv np(1 - p)$$

Además, se ha demostrado que la distribución binomial es asimétrica cuando $p \neq 0.5$.

Cabe señalar que el caso de asimetría positiva es relevante para los gráficos de control p y np , los cuales están basados en la distribución binomial, esto debido a que en control de procesos los valores de fracción de no conformes (p) son siempre mucho más bajos que 0.5.

En la siguiente sección, y debido a su importancia, se desarrolla la comprobación de la aproximación normal a la distribución binomial.

4.1.3 Aproximación de la distribución binomial a la normal

4.1.3.1 Características de la distribución normal

Se dice que una variable aleatoria X sigue una distribución normal $X \sim N(\mu, \sigma^2)$, con media (μ) y varianza (σ^2), si su función de densidad es igual a la ecuación (d.1).

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (\text{d.1})$$

Dónde:

$$E(X) = \mu$$

$$\text{Var}(X) = \sigma^2$$

Una característica importante de la distribución normal, es que es simétrica con respecto a μ , por lo que $P[X \leq \mu] = P[X \geq \mu] = 1/2$. Por esta razón la mediana y la moda coinciden con la media (μ). La Figura 11 muestra la forma de esta distribución.

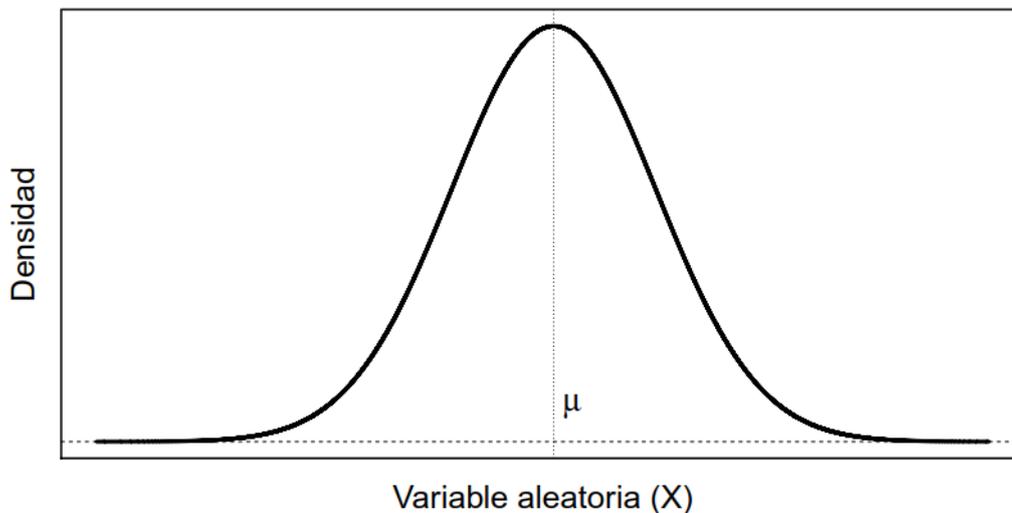


Figura 11. Forma de la distribución normal o gaussiana

4.1.3.2 Comprobación de la aproximación normal a la distribución binomial

Se conoce que la aproximación de Sterling para la función factorial $a!$ es igual a la expresión dada en (d.2):

$$a! = a^a e^{-a} \sqrt{2\pi a} \quad (\text{d.2})$$

Usando (d.2) para las funciones factoriales $n!$, $x!$ y $(x - n)!$ de la ecuación (a.1) se obtienen (d.3), (d.4) y (d.5):

$$n! = n^n e^{-n} \sqrt{2\pi n} \quad (\text{d.3})$$

$$x! = x^x e^{-x} \sqrt{2\pi x} \quad (\text{d.4})$$

$$(x - n)! = (x - n)^{(x-n)} e^{-(x-n)} \sqrt{2\pi(x-n)} \quad (d.5)$$

Reemplazando (d.3), (d.4) y (d.5) en (a.1) se obtiene (d.6) y consecuentemente (d.7):

$$P(x) = \frac{n^n e^{-n} \sqrt{2\pi n}}{[x^x e^{-x} \sqrt{2\pi x}] [(x-n)^{(x-n)} e^{-(x-n)} \sqrt{2\pi(x-n)}]} p^x q^{n-x} \quad (d.6)$$

$$\begin{aligned} &= \left(\frac{p}{x}\right)^x \left(\frac{q}{n-x}\right)^{n-x} n^n \sqrt{\frac{n}{2\pi x(n-x)}} \\ &= \left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x} \sqrt{\frac{n}{2\pi x(n-x)}} \end{aligned} \quad (d.7)$$

Trabajando solo con los logaritmos naturales de los dos primeros factores de (d.7) y definiendo $\delta = x - np$, se obtiene que:

$$\ln\left(\frac{np}{x}\right) = -\ln\left(1 + \frac{\delta}{np}\right) \quad \text{y} \quad \ln\left(\frac{np}{n-x}\right) = -\ln\left(1 - \frac{\delta}{np}\right)$$

Y dado que:

$$\ln\left(\left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x}\right) = x \ln\left(\frac{np}{x}\right) + (n-x) \ln\left(\frac{np}{n-x}\right)$$

Entonces:

$$\ln\left(\left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x}\right) = -x \ln\left(1 + \frac{\delta}{np}\right) - (n-x) \ln\left(1 - \frac{\delta}{np}\right)$$

Usando la expansión $\ln(1+a) = a - \frac{1}{2}a^2 + \mathcal{O}(a^3)$ en la anterior expresión se llega a obtener (d.8)

$$\ln\left(\left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x}\right) = -\frac{\delta^2}{2npq} + \mathcal{O}\left(\frac{\delta^3}{n^2}\right) \quad (d.8)$$

En (d.8) el error de truncamiento (\mathcal{O}) se vuelve insignificante para valores altos de n por lo que puede ser omitido para estos casos, entonces de (d.8) y asumiendo que n es grande, se obtiene (d.9):

$$\left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x} = e^{-\frac{\delta^2}{2npq}} = e^{-\frac{1}{2} \frac{(x-np)^2}{npq}} \quad (d.9)$$

Ahora dé $\delta = x - np$ se puede obtener $x = np + \delta$ y $n - x = nq - \delta$. Reemplazando estas igualdades en el tercer término de (d.7) se obtiene (d.10):

$$\begin{aligned}\sqrt{\frac{n}{2\pi x(n-x)}} &= \sqrt{\frac{n}{2\pi(np+\delta)(nq-\delta)}} \\ &= \sqrt{\frac{1}{2\pi n(p+\frac{\delta}{n})(q-\frac{\delta}{n})}}\end{aligned}\quad (d.10)$$

En (d.10) el factor δ/n se vuelve insignificante para valores grandes de n por lo que en este caso (d.10) puede ser simplificado y tomar la forma mostrada en (d.11):

$$\sqrt{\frac{n}{2\pi x(n-x)}} = \sqrt{\frac{1}{2\pi(npq)}}\quad (d.11)$$

Reemplazando (d.9) y (d.11) en (d.7) se obtiene que la función de probabilidad binomial (a.1) puede ser aproximada por la ecuación (d.12), con valores de muestra (n) grandes o equivalentemente también con valores de np y npq grandes.

$$P(x) \approx \sqrt{\frac{1}{2\pi(npq)}} e^{-\frac{1}{2} \frac{(x-np)^2}{npq}}\quad (d.12)$$

Dónde:

$$E(X) = np$$

$$Var(X) = npq$$

Comparando (d.12) con (d.1) se puede apreciar que existe una similitud entre estas dos expresiones, de lo que se puede concluir que la función de distribución binomial (a.1) puede ser aproximada por la distribución normal (d.1), con parámetros $\mu = np$ y $\sigma = \sqrt{npq}$, para valores grandes de n , np o npq , tal como se muestra en (d.13)

$$P(x) = \theta(x) \approx \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}\quad (d.13)$$

Dónde:

- $x = \text{variable aleatoria}$
- $\mu = np$
- $\sigma = \sqrt{npq}$

Aplicando a (d.13) el método de momentos para una variable continua, es posible verificar que las expresiones matemáticas para μ y σ^2 son:

$$\begin{aligned}
 E(x) &= \int_{-\infty}^{\infty} x \theta(x) dx = np \\
 Var(x) &= E(x^2) - [E(x)]^2 \\
 &= \int_{-\infty}^{\infty} x^2 \theta(x) dx - \left[\int_{-\infty}^{\infty} x \theta(x) dx \right]^2 \\
 &= npq
 \end{aligned}$$

4.1.4 Principios de gráficos de control np y p

Vamos a suponer que un proceso está bajo control estadístico y que este genera 10% de productos no-conformes, o en otras palabras, que tiene una fracción de no-conformes igual a $p = 0.1$. Si de ese proceso se tomara una muestra de tamaño $n = 1$, entonces la probabilidad de que n fuese no-conforme sería igual a $p = 0.1$ y de que fuese conforme $(1 - p) = q = 0.9$. Sumando estas dos probabilidades se obtiene:

$$p + q = 0.1 + 0.9 = 1$$

Como se puede apreciar, la suma de las probabilidades es igual a uno, esto es lógico ya que la sumatoria de las probabilidades de todos los eventos posibles, debe ser igual a uno. Y dado que en el ejemplo solo existen dos posibilidades (no-conforme y conforme) entonces $0.1 + 0.9 = 1$.

Ahora bien, si $n=2$ entonces se tienen las siguientes probabilidades:

<i>Posibilidades</i>	<i>Probabilidades</i>	<i>Valor de la probabilidad</i>
Dos no-conformes	$p \cdot p = p^2$	0.01
Un no-conforme	$pq + qp = 2pq$	0.18
Cero no-conformes	$q \cdot q = q^2$	0.81
Total:		1

Es fácil reconocer que la suma de las probabilidades $p^2 + 2pq + q^2$ es una expansión de $(p + q)^2$. Si se tuviese que $n = 3$, entonces la suma de las probabilidades serían equivalentes a una expansión de $(p + q)^3$, para $n = 4$ sería $(p + q)^4$, y así sucesivamente. De estos ejemplos se puede deducir que las probabilidades, para cualquier valor de n , pueden ser calculadas mediante la expresión binomial $(p + q)^n = 1$, por lo que la probabilidad de que se encuentren un x número de productos no-conformes en la muestra, puede ser calculada mediante la ecuación (3), que es la función de probabilidad binomial.

$$P(x) = Q(n, x)p^x q^{n-x}, \quad (3)$$

$$0 \leq p \leq 1,$$

$$x = 0, 1, 2 \dots n$$

Dónde: $Q(n, x) = \binom{n}{x} = \frac{n!}{x!(n-x)!}$

$$q = 1 - p ; n = \textit{tamaño de muestra}$$

De lo expuesto en los párrafos anteriores, es posible deducir que se podría determinar un número límite de no-conformes por muestra y mediante la ecuación (3) determinar la posibilidad de que este sea excedido, dando lugar así a los principios fundamentales de los GCs p y np .

Para poder aplicar la teoría de probabilidad binomial antes descrita para construir GCs, se debe primeramente determinar la fracción de no-conformes (p) del proceso en cuestión. Esto se lo puede hacer tomando una cantidad razonable de muestras, digamos 80, durante un periodo en que el proceso esté operando bajo control estadístico, de forma que cualquier producto generado por el proceso, tenga una probabilidad p de ser no-conforme.

Un ejemplo de resultados de muestreo para $n=100$ se da en Tabla 2, mientras que la Figura 12 muestra un histograma de ellos.

Tabla 2: Ejemplo de resultados de muestreo para determinar p , utilizando $n = 100$.

Número de no-conformes	Frecuencias
6	3
7	6
8	9
9	14
10	17
11	12
12	7
13	6
14	3
15	2
16	1

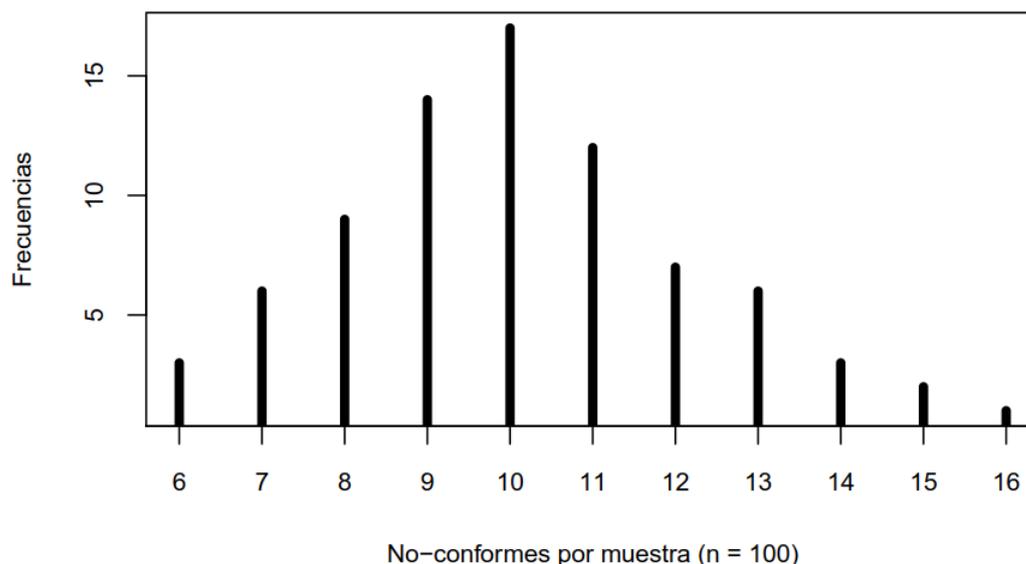


Figura 12. Histograma de los resultados del muestreo hecho para calcular p

El número promedio de no-conformes por muestra (np), puede ser estimado dividiendo el número total de no-conformes por el número total de muestras, en este caso $np = 810/80 = 10.1$, de donde se obtiene que $p = 10.1/100 \approx 0.1$.

4.1.4.1 Ecuaciones para computar los límites de control de los gráficos de control np y p

Para construir los gráficos de control (GCs) np y p del tipo Shewhart, que es el tema de estudio de esta sección, se asume que la distribución binomial es aproximadamente normal, la idoneidad de la aproximación habitualmente se comprueba mediante reglas empíricas como $np \geq 5$ (o 10). Por ejemplo, para el caso expuesto en la sección anterior, se tiene que $np = 100 * 0.1 = 10$ por lo que en este caso se concluye que la aproximación es buena y por consiguiente se asume que la desviación estándar del número de no-conformes por muestra es $\sigma_x = \sqrt{npq}$, o también que la desviación estándar de la fracción de no-conformes es $\sigma_{(x/n)} = \sqrt{pq/n}$. Los límites de control se los calcula mediante las ecuaciones incluidas en la Tabla 3, cabe señalar que estas han sido derivadas con base en la presunción de que la aproximación normal a la distribución de binomial es adecuada, ver (Mitra, 2008, pág. 374) y (Montgomery, 2009, pág. 288).

Tabla 3: Ecuaciones para calcular los límites de control para los graficos np y p tipo Shewhart para un tamaño de muestra n

<i>Gráfico de control np</i>	<i>Gráfico de control p</i>
<i>Línea central = np</i>	<i>Línea central = p</i>

$LCS = np + (K_{\alpha/2}) \sigma_x$ $LCI = np - (K_{\alpha/2}) \sigma_x$ <p>Dónde:</p> <ul style="list-style-type: none"> • $\sigma_x = \sqrt{npq}$ • $(K_{\alpha/2}) = 3$ • $LCS = \text{Limite de control superior}$ • $LCI = \text{Limite de control inferior}$ 	$LCS = p + (K_{\alpha/2}) \sigma_{(x/n)}$ $LCI = p - (K_{\alpha/2}) \sigma_{(x/n)}$ <p>Dónde:</p> <ul style="list-style-type: none"> • $\sigma_{(x/n)} = \sqrt{pq/n}$ • $(K_{\alpha/2}) = 3$
---	--

Como se puede apreciar en la Tabla 3, los límites de control se los ubica a una distancia de tres variaciones estándar ($\pm 3\sigma$) de la media. El uso de los límites “3-sigma” fueron propuestos originalmente por Walter A. Shewhart, las razones del porqué él eligió esos límites los detalla en su extraordinario libro “Statistical Method From the Viewpoint of Quality Control (Shewhart, 1939). Gurús de la calidad como Deming (1989) y Juran (1990) han validado el uso de los límites “3-sigma” y actualmente son universalmente aceptados.

4.1.4.2 Relación entre prueba de hipótesis y los gráficos de control por atributos

Aunque en esta tesis no se ahonda en las razones para usar los límites 3-sigma en GCs por atributos, algo que sí es importante analizar es la significancia de esos límites en términos de la probabilidad de los errores tipo-I y tipo-II para la hipótesis H_0 , que se detalla en la Tabla 4, mediante la cual se determina si el proceso está o no bajo control.

Tabla 4: Relación entre prueba de hipótesis y los GCs por atributos

$H_0: E(x_i) = \mu_0 \rightarrow$ Si se acepta H_0 se concluye que el proceso está bajo control $H_1: E(x_i) \neq \mu_0 \rightarrow$ Si se acepta H_1 se concluye que el proceso no está bajo control Dónde: <ul style="list-style-type: none"> • $x_i = \text{Valor que toma el estadístico } X \text{ en la } i\text{-ésima muestra}$ • <ul style="list-style-type: none"> – Para el GC np: $X = \text{número de no-conformes por muestra}$ – Para el GC p: $X = \text{fracción de no-conformes por muestra}$ • $\mu_0 = \text{Valor promedio de } X$ 	
Error de tipo I:	Este tipo de error consiste en rechazar H_0 cuando la hipótesis es verdadera. La probabilidad de cometer este error se lo denota con la letra griega alfa (α). En otras palabras, α es la probabilidad de que el GC indique que el proceso está fuera de control (por medio de un punto fuera de los límites) cuando en verdad no lo está.

Error de tipo II:	<p>Es el error que consiste en no rechazar H_0 cuando la hipótesis es falsa. La probabilidad de cometer este error se lo denota con la letra griega beta (β).</p> <p>En otras palabras, β es la probabilidad de que el GC no identifique que el proceso está fuera de control cuando en verdad si lo está.</p>
-------------------	---

De la Tabla 4 se puede apreciar que si se acepta H_0 y rechaza H_1 , entonces se asume que el proceso está bajo control. Por el contrario, si se rechaza H_0 y acepta H_1 , entonces se asume que el proceso está fuera de control. Estas hipótesis son comprobadas cada vez que se traza el valor del estadístico X en el gráfico de control, ya que si X cae dentro de los límites de control se acepta H_0 y se rechaza H_1 .

Ahora bien, cuando se utiliza un GC por atributos el tipo de prueba de hipótesis que se lleva a cabo es bilateral, o de dos colas, y antes de realizarla es necesario definir un valor para la probabilidad de cometer un error de tipo I para cuando el proceso está bajo control, conocido también como nivel de significación alfa (α). Asumiendo que la v.a.³ X tiene una distribución normal y dado que esta es simétrica, ver Figura 1, se puede obtener α sumando las áreas de las colas que están fuera de los límites 3-sigma, el resultado es $\alpha = 0.0027$. Esta probabilidad puede ser calculada mediante la ecuación (4) usando μ y σ de la Tabla 3 según sea el caso.

$$\alpha_I + \alpha_S = \alpha = P(LCI > X > LCS) = \left[1 - \int_{LCI=\mu-3\sigma}^{LCS=\mu+3\sigma} f(x) dx \right] \approx 0.0027 \quad (4)$$

Dónde:

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

- $\alpha_I =$ Area de probabilidad de la cola Inferior $P(X < LCI)$
- $\alpha_S =$ Area de probabilidad de la cola superior $P(X > LCS)$
- $\alpha_I = \alpha_S \approx 0.00135$

El valor $\alpha = 0.0027$ puede ser visto como la probabilidad conjunta de que la variable X caiga fuera de cualquiera de los límites de control cuando el proceso está bajo control. Ahora bien, dado que las colas alfa, en teoría, son iguales se tiene que $\alpha_I = \alpha_S = \alpha/2 \approx 0.00135$, que es la probabilidad de que X caiga fuera de uno de los límites (LCI o LCS) cuando el proceso está bajo control. Cabe destacar que en esta tesis se utiliza el término **colas alfa** para referirse a α_I y α_S .

³ v.a = variable aleatoria

Otra forma de ver estas probabilidades es que cuando el proceso está bajo control, y ya que $1 - \alpha = 1 - 0.0027 = 0.9973$, hay un 99.73% de probabilidad que los puntos trazados en el gráfico caigan dentro de los límites de control.

4.1.4.3 Variaciones del proceso y como se manifiestan en el gráfico de control

La Figura 13 y la Figura 14 muestran ejemplos de gráficos de control (GCs) np y p con líneas centrales y límites basados en los μ y σ calculados usando los datos de la Tabla 2. Los puntos trazados en estos gráficos pertenecen a 20 nuevas muestras, se puede observar que todos los puntos caen dentro de los límites de control, por lo que se concluye que el proceso está bajo control.

Ryan (2011, pág. 184) menciona que, desde un punto de vista estadístico, los GCs np y p son equivalentes y que no hay necesidad de usar los dos gráficos al mismo tiempo, además dice que es razonable utilizar el GC np si el tamaño de muestra es constante. Por otro lado, Montgomery (2009, pág. 301), menciona que para personal sin formación en estadística el gráfico np es más fácil de interpretar.

Para desarrollar la investigación de GCs basados en la distribución binomial se optó por el GC p asumiendo que n es constante, aunque los resultados, como se explica en la sección 5.5.3.6, también aplican al caso de n variable. Por esta razón de este punto en adelante, se hablará solo acerca de GCs p .

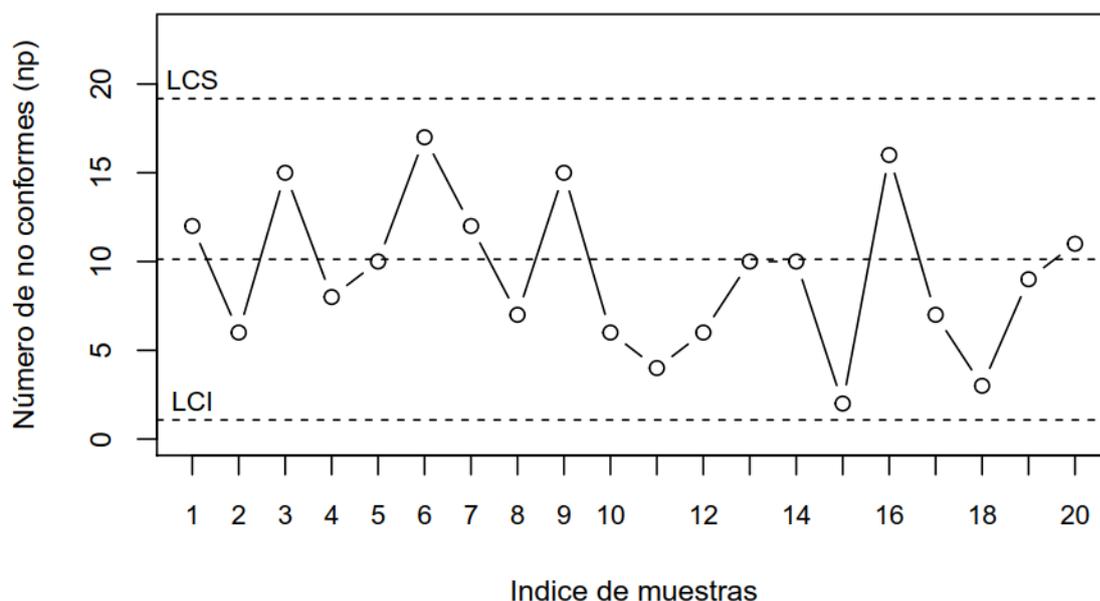


Figura 13. Ejemplo de un GC np tipo Shewhart

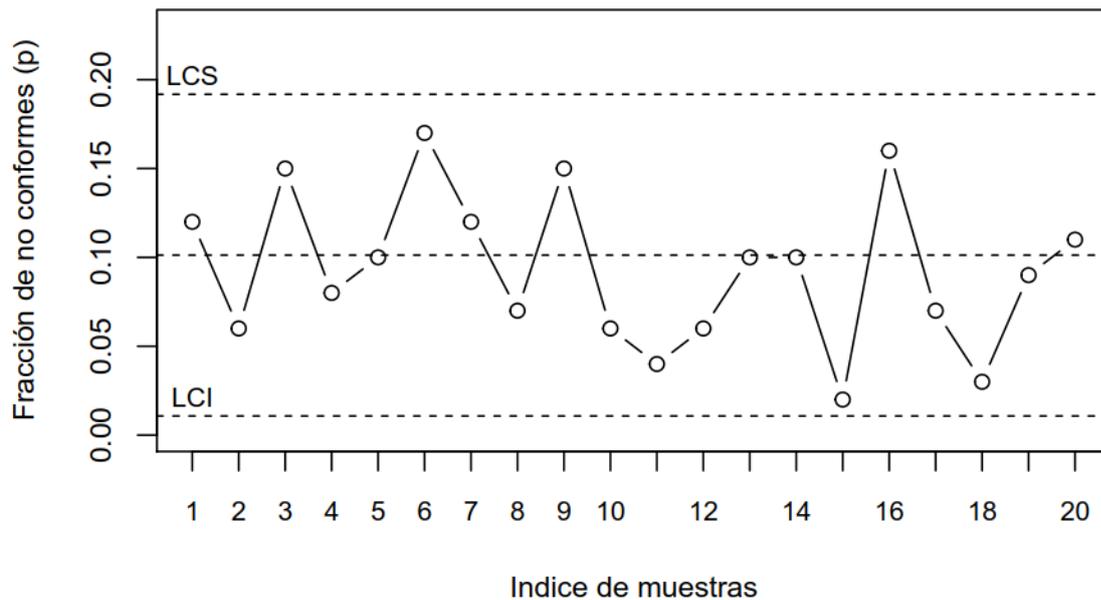
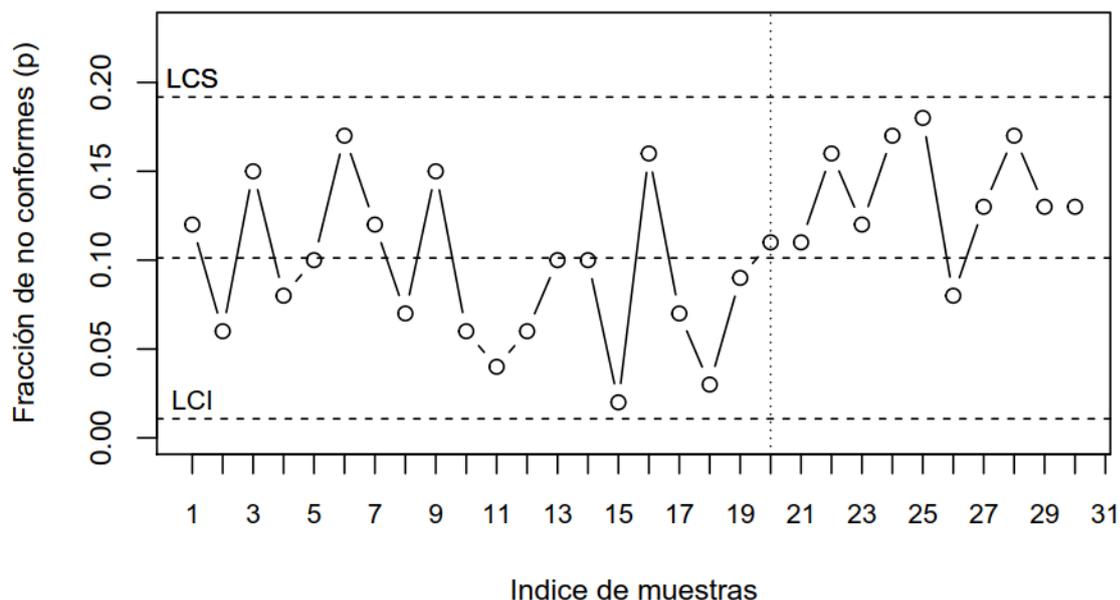
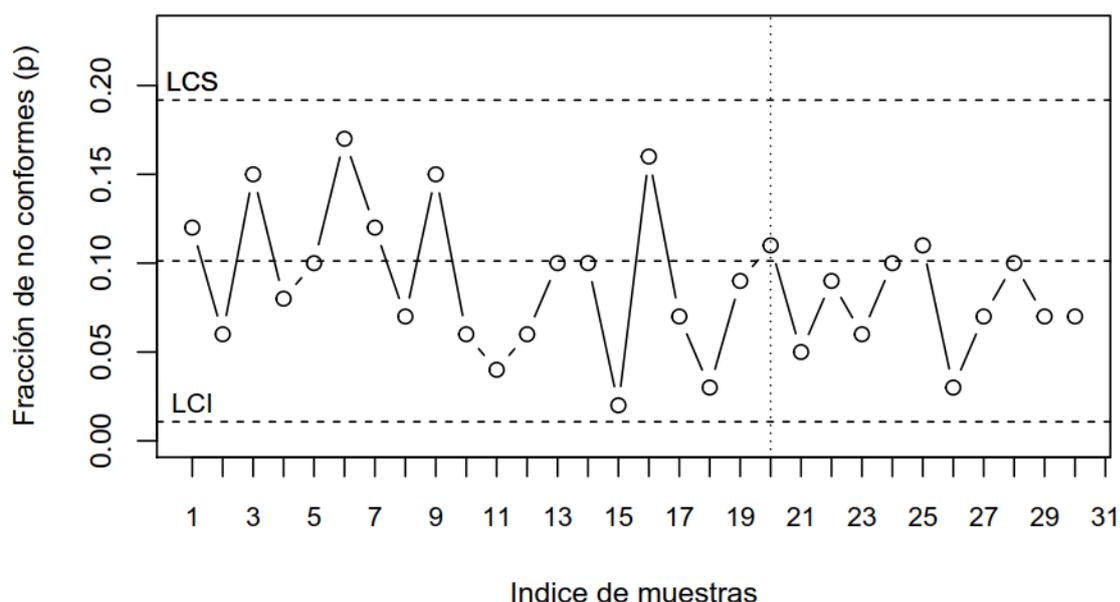


Figura 14. Ejemplo de un GC p tipo Shewhart

La Figura 15 y la Figura 16 muestran dos escenarios muy importantes para el control de procesos con GCs por atributos. En la Figura 15, los últimos 10 puntos vienen de un incremento del 30% en la fracción de no-conformes (p), es decir $p = 0.1 * 1.3 = 0.13$; incrementos de este tipo son indeseables ya que manifiestan un deterioro del proceso. Como se puede observar, en este caso la mayoría de puntos están por encima de la línea central. Por otro lado, en la Figura 16, los últimos 10 puntos vienen de una reducción del 30% en p , es decir $p = 0.1 * 0.7 = 0.07$; reducciones de este tipo son deseables ya que indican una mejora del proceso. Como se puede apreciar en este caso la mayoría de puntos están por debajo de la línea central.

Determinar la potencia o capacidad de un GC para detectar mejoras o deterioros del proceso, como los expuestos en el párrafo anterior, es de suma importancia. En la siguiente sección se describe el parámetro más comúnmente utilizado para cuantificar la capacidad de monitorización (también llamado potencia de detección) de GCs por atributos y que fue empleado a lo largo del trabajo de investigación resumido en este documento.

Figura 15. Ejemplo de deterioro del proceso (incremento de p)Figura 16. Ejemplo de mejora del proceso (reducción de p)

4.1.5 El ARL: Un parámetro para evaluar la capacidad de los gráficos de control por atributos para detectar cambios en el proceso

Como se ha explicado anteriormente, cuando uno o más puntos caen fuera de los límites de control de un GC por atributos, se asume que el proceso monitorizado está fuera de control.

El ARL (o Average Run Length en inglés) es el promedio de puntos que caen dentro de los límites de control antes de que uno caiga fuera de ellos, también podría ser visto como el promedio de muestras requeridas para detectar un cambio de cierta magnitud en el proceso. El

ARL es un parámetro ampliamente empleado para cuantificar y evaluar la capacidad de monitorización de gráficos de control por atributos, y fue el parámetro que se utilizó para analizar la capacidad de monitorización de todos los GCs incluidos en esta investigación. Por su importancia a continuación se deriva la expresión matemática utilizada para computar su valor.

Vamos a suponer que ω denota la probabilidad de que un punto caiga fuera de cualquiera los límites de control (LCI o LCS). En este caso, la probabilidad de que el punto caiga fuera de los límites en la primera muestra es igual a ω , o de que caiga fuera no en la primera sino en la segunda muestra es igual a $(1 - \omega)\omega$, o de que caiga fuera no en la primera ni en la segunda, sino en la tercera muestra es igual a $(1 - \omega)^2\omega$, y así sucesivamente. De lo anterior es posible deducir que el ARL es igual a la serie infinita definida por la ecuación (5), ver Mitra. (2008, pág. 277)

$$ARL = \sum_{j=1}^{\infty} j(1 - \omega)^{j-1} \omega \quad (5)$$

$$ARL = \omega \sum_{j=1}^{\infty} j(1 - \omega)^{j-1} \quad (6)$$

De (5) se obtiene (6) y dado que la serie sumatoria infinita se la puede obtener mediante $1/[1 - (1 - \omega)]^2$, realizando las respectivas simplificaciones, se obtiene que el ARL es igual a (7).

$$ARL = \frac{\omega}{[1 - (1 - \omega)]^2} = \frac{1}{\omega} \quad (7)$$

4.1.5.1 ARL para el gráfico de control p cuando el proceso está bajo control (ARL_0)

Cuando el proceso está bajo control el ω de ecuación (7) es igual a la probabilidad de cometer un error de Tipo I (α), de esto se obtiene la ecuación (8).

$$ARL_0 = \frac{1}{\omega} = \frac{1}{\alpha} \quad (8)$$

El ARL_0 también puede ser calculado mediante la ecuación (9) debido a que la probabilidad de error tipo I (α), es igual a la suma de las probabilidades de las colas alfa. Cabe recordar que en la sección 4.1.4.2 se definió el término “colas alfa” el cual es usado a lo largo de esta tesis

para referirse a las probabilidades de que, en un gráfico de control, un punto caiga fuera de uno de los límites (LCI o LCS) cuando el proceso está bajo control.

El ARL_0 representa la media de puntos trazados dentro de los límites de control antes de que un punto caiga fuera de ellos. En este caso, y ya que el proceso está bajo control, el punto que cae fuera de los límites es considerado como una falsa alarma.

$$ARL_0 = \frac{1}{\alpha} = \frac{1}{\alpha_I + \alpha_S} \quad (9)$$

Dónde:

$\alpha_S = \text{Cola alfa superior}$ (Probabilidad de que un punto caiga fuera del límite de control superior o LCS, cuando el proceso está bajo control)

$\alpha_I = \text{Cola alfa inferior}$ (Probabilidad de que un punto caiga fuera del límite de control inferior o LCI, cuando el proceso está bajo control)

Para el caso del Gráfico de Control p , las áreas de las colas alfa se las calcula utilizando la probabilidad binomial acumulada, como se muestra en las ecuaciones (10) y (11).

$$\alpha_S = P\{x > nLCS|_p\} = 1 - \sum_{x=0}^{x=nLCS} \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1,2 \dots nLCS \quad (10)$$

$$\alpha_I = P\{x < nLCI|_p\} = \sum_{x=0}^{x=nLCI} \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0,1,2 \dots nLCI \quad (11)$$

Dónde:

p = Fracción de no-conformes del proceso

n = Tamaño de muestra

4.1.5.2 ARL para el grafico de control p cuando el proceso está fuera de control (ARL_1)

Cuando el proceso ha sufrido un cambio, es decir el valor de p ha sido modificado debido a un incremento o disminución en la fracción de no-conformes del proceso, se tiene que el ω de la ecuación (7) es igual a $1 - \beta$, donde β es la probabilidad de cometer un error de Tipo II, por lo que el ARL para este caso, y que en esta tesis se denota como ARL_1 , está dado por la ecuación (12).

El ARL_1 representa la media de puntos trazados dentro de los límites de control, antes de que uno caiga fuera de ellos, cuando el proceso está fuera de control.

$$ARL_1 = \frac{1}{1 - \beta} \quad (12)$$

Dónde:

$$\beta = Pr\{x \leq nUCL|_{p_1}\} - Pr\{x \leq nLCL|_{p_1}\}$$

$p_1 = p$ -modificado

p_1 podría ser: $p_1 < p$ o $p_1 > p$

$Pr =$ Probabilidad binomial acumulada

4.1.6 Gráfico de Control p Shewhart: Colas alfa y ARL_0

En la sección 4.1.4.1, específicamente en la Tabla 3, se presentó la ecuación que sirve para calcular los límites de control del GC p Shewhart y se mencionó que esa ecuación fue derivada bajo la presunción de que la aproximación normal a la distribución binomial es adecuada. En esa sección también se mencionó la regla empírica $np \geq 5$ (o 10), la cual es comúnmente utilizada para verificar si existen las condiciones para una buena aproximación. Por otra parte, en la sección 4.1.4.2 también se definió el término **colas alfa**, que hace referencia a las probabilidades de que, en un gráfico de control, un punto caiga fuera de uno los límites cuando el proceso está bajo control.

Ahora bien, cuando la aproximación normal a la distribución binomial es buena, lo habitual es simplemente asumir que el área de cada una de las colas alfa tiene un valor, sino igual, al menos muy cercano a 0.00135, ver ecuación (4). En consecuencia, y usando la ecuación (9), se obtiene que $ARL_0 = 1/(\alpha_I + \alpha_S) = 1/(0.00135 + 0.00135) = 1/(0.0027) \approx 370$. Esto significa que cuando el proceso está bajo control, en promedio, 370 puntos caerían dentro de los límites de control antes de que, por azar, uno cayese fuera de ellos. En este caso, y debido a que el proceso está bajo control, los puntos que caen fuera de los límites serían considerados como falsas alarmas.

Sin embargo, y como se demuestra más adelante, la presunción de que las colas alfa son iguales a 0.00135 es errónea y por lo tanto asumir que $ARL_0 = 370$ también lo es. Esto se debe principalmente a que cuando $p \ll 0.5$ la distribución binomial es altamente asimétrica, lo que causa que el ajuste de la aproximación normal sea muy pobre en las colas de la distribución binomial.

En las siguientes secciones se analiza el comportamiento real de las colas alfa y también su efecto en el ARL_0 del GC p .

4.1.6.1 Comportamiento real de las colas alfa para el gráfico de control p Shewhart

Regresemos por un momento a la sección 4.1.4 en donde está la Tabla 2 que contiene los resultados de un muestreo hecho con $n = 100$. De esos datos se obtuvo que la fracción de no-conformes era aproximadamente $p = 0.10$ y en base a eso se construyó el GC mostrado en la Figura 14 con límites LCI = 0.01 y UCL = 0.19. En este caso se obtiene que $np = (100 \cdot 0.1) = 10$, que cumple con la regla empírica $np \geq 10$, por lo que se asumiría que la aproximación normal a la binomial es adecuada y que en consecuencia $\alpha_I = \alpha_S = 0.00135$ y que el ARL_0 del GC es 370. Sin embargo utilizando las ecuaciones (10) y (11), mediante las cuales se puede calcular los verdaderos valores de las colas alfa, se obtiene que $\alpha_I = 0.00032$ y $\alpha_S = 0.00198$ por lo que el verdadero ARL_0 es 435. Como se puede apreciar los verdaderos valores de las colas alfa no son iguales a 0.00135, además llama mucho la atención la gran diferencia que existe entre la cola alfa inferior (α_I) y el valor nominal 0.00135.

Cabe mencionar que el hecho de que las colas alfa en los GCs Shewhart no alcanzan a obtener el valor nominal 0.00135, era algo que ya había sido identificado previamente. Un ejemplo de esto es el trabajo publicado por Ryan (2011), quien en la página 190 de su libro, muestra una tabla con algunos valores reales de colas alfa que difieren mucho del valor nominal. Sin embargo, luego de estudiar lo publicado por Ryan y otros autores con respecto a las colas alfa, se concluyó que la información existente no daba una respuesta definitiva respecto al comportamiento de las colas alfa y que se requería de un análisis más profundo. Por ello se decidió caracterizar el comportamiento de las colas de una forma detallada y además presentar los resultados mediante un método de fácil interpretación.

Para lograr este objetivo se creó un algoritmo en R^4 capaz de computar las colas alfa de cualquier combinación p y n , y además se escogió el método gráfico para presentar y analizar los resultados. La caracterización fue realizada para $p = 0.01(0.01)0.20$ (notación que indica que los cálculos fueron realizados desde $p = 0.01$ hasta $p = 0.20$, en incrementos de 0.01) usando rangos de tamaños de muestra específicos para cada p , y cuyos valores mínimo y máximo fueron determinados mediante la ecuación (13). Esta ecuación es comúnmente utilizada para determinar el tamaño de muestra para el GC p , ver (Duncan, 2000, pág. 451).

$$n = 9p(1 - p)/d^2 \quad (13)$$

⁴ R es un programa computacional enfocado a la estadística

Ejemplos de los resultados obtenidos se exponen en la Figura 17 y también en las figuras del Anexo 1. Estas figuras muestran los verdaderos valores de las colas alfa para $p = 0.1, 0.05$ y 0.01 dentro de un rango de n específicos para cada p . Para establecer los valores mínimo y máximo del rango de tamaños de muestras⁵ (o rango de n), se aplicaron los siguientes criterios: i) que con todas las combinaciones de n y p se debían obtener límites de control inferior (LCI) y superior (LCS) y ii) que el tamaño de muestra debía ser lo suficientemente grande. Para cumplir estos criterios se utilizó en (13) $d = 0.99p$ para obtener el valor de n mínimo y $d = p/2$ para el n máximo. En los gráficos el rango de tamaños de muestra se denota como: n rango = n mínimo (1) n máximo, por ejemplo en la Figura 17 se muestra n rango = 83(1)324, indicando que los cálculos fueron realizados desde el tamaño de muestra 83 hasta el 324 en incrementos de una unidad.

En este punto cabe aclarar la diferencia que existe entre los conceptos “rango de tamaños de muestra” y “tamaño de muestra variable”. El primero hace referencia a una secuencia de valores de n que tiene un mínimo y máximo y cuyo valor se incrementa acorde a una magnitud específica pre-establecida, por ejemplo 83(1)324; mientras que el segundo se refiere a un número no específico de muestras n_i ($i = 1, 2, 3 \dots$) cuyos tamaños pueden variar de manera arbitraria, por ejemplo $n_1 = 95, n_2 = 125, n_3 = 87$ etc. El concepto de “tamaño de muestra variable” es utilizado para el caso de monitorización de procesos mediante muestras no constantes, tema al cual se hace alusión en este documento solo en las secciones 5.5.3.6 y 5.6.3.6. Mientras que “rangos de tamaños de muestra” es utilizado a lo largo de esta tesis para analizar una serie, o grupo, de gráficos individuales ordenados en una secuencia continúa. Por ejemplo, para el caso de $p = 0.1$ y n rango = 83(1)324 de la Figura 17, se tienen un grupo de 240 gráficos, es decir un gráfico para cada combinación de p y n .

Regresando a los resultados obtenidos, en ellos se puede apreciar que las colas alfa de los gráficos de control construidos con el GC p Shewhart no tienen un valor constante, sino que estas más bien oscilan por encima (cola superior) y por debajo (cola inferior) de 0.00135, y que además la cola alfa inferior se acerca a cero a medida que el tamaño de muestra se reduce. Nótese que inclusive con tamaños de muestras grandes las colas alfa no alcanzan el valor de 0.00135. Por otra parte, de los resultados también es posible deducir que los valores de las colas alfa varían en función de los parámetros p y n . El comportamiento de las colas alfa fue similar para todos los valores de p utilizados en el estudio.

⁵ Rango de tamaños de muestras = rango de n

Nótese que si se aplicara la regla empírica $np \geq 5$ (o 10) a cualquiera de las combinaciones p y n de la Figura 17, nos llevaría a la errónea conclusión de que la aproximación normal a la binomial es adecuada y que en consecuencia cada cola alfa tendría un valor, sino igual, al menos muy cercano a 0.00135. Por esta razón se concluye que esta regla empírica no es adecuada para establecer la idoneidad de las colas alfa del GC p Shewhart.

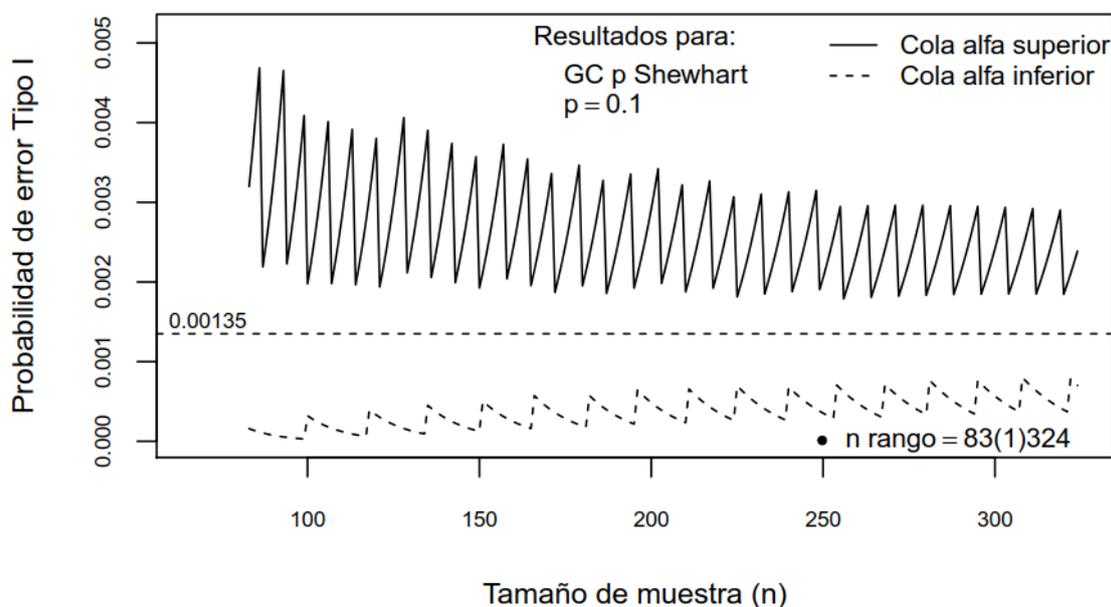


Figura 17. GC p Shewhart - Comportamiento real de la colas alfa para $p = 0.1$

4.1.6.2 Comportamiento real del ARL_0 para el gráfico de control p Shewhart

La Figura 18 muestra los ARL_0 que se obtienen con las colas alfa de la Figura 17 y que fueron computados mediante la ecuación (9). Como se puede apreciar, los ARL_0 no tienen un valor constante como comunmente se supone, sino que más bien estos varían y fluctúan alrededor de 370. Además, nótese que la mayoría de los valores ARL_0 están por debajo de 370. Resultados del ARL_0 para $p = 0.05$ y 0.01 se incluyen en el Anexo 1.

Cabe destacar que el valor del ARL_0 es de suma importancia para poder realizar una efectiva monitorización del proceso. Si el ARL_0 es muy bajo las falsas alarmas (o puntos que caen fuera de los límites cuando el proceso está bajo control) ocurrirían con mucha frecuencia. En cambio un ARL_0 muy alto podría perjudicar la rapidez del gráfico para detectar cambios en el proceso.

Los resultados demuestran que la variación y fluctuación de las colas alfa afecta directamente al ARL_0 y que su valor real puede estar muy alejado de 370. Además, debido a que las colas alfa están en función de p y n , es lógico deducir que el ARL_0 también lo está. Debido a la importancia del ARL_0 , y dado que su valor varía en función de p y n , lo más adecuado sería

computar su valor real y determinar si este es o no aceptable. Sin embargo, para hacer esto primero hay que establecer un rango de valores ARL_0 que pudiesen ser considerados aceptables. En esta tesis el rango de ARL_0 aceptables que se utiliza es 250 a 450, el cómo se llegó a determinar este rango se detalla en la sección 5.1.3.

Nótese en la Figura 18 que el 91% de los gráficos p Shewhart contruidos con $p = 0.1$ y $n = 83(1)324$ tendrían un ARL_0 aceptable según el criterio establecido en esta tesis.

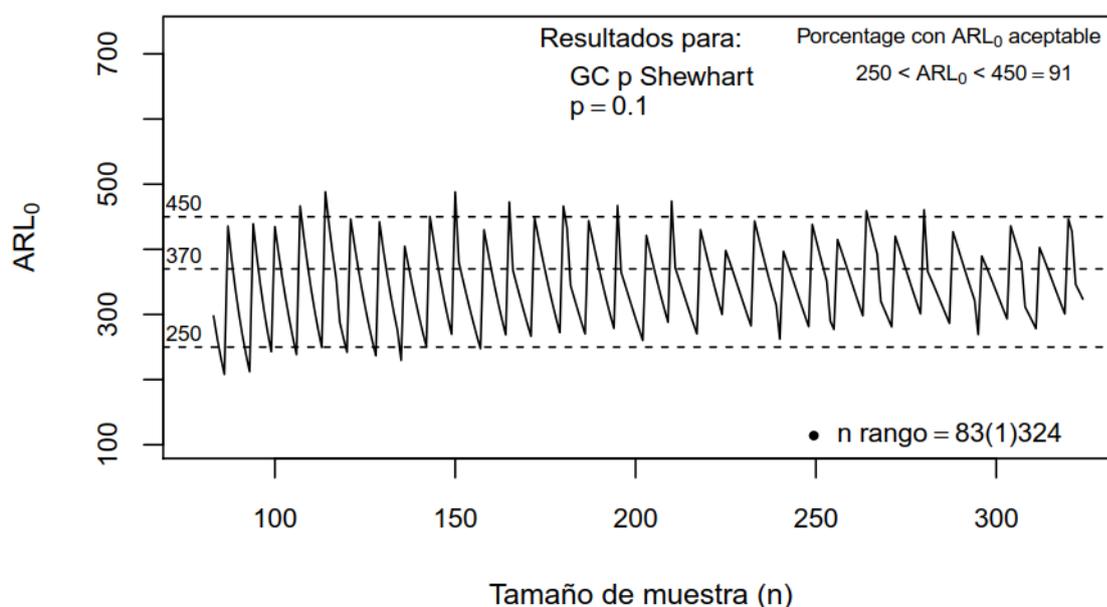


Figura 18. GC p Shewhart – Fluctuación del ARL_0 para $p = 0.1$

4.1.7 Análisis del ajuste de la aproximación normal a la distribución binomial cuando $p \ll 0.5$

En las secciones previas se demostró que las colas alfa del GC p Shewhart no tienen un valor constante sino que más bien su valor fluctúa en función de p y n . En esta sección se analiza el por qué las colas alfa no llegan a obtener el valor deseado de 0.00135.

Tomando el ejemplo expuesto en la Tabla 2, la Figura 19 muestra la probabilidad binomial para $n = 100$ y $p = 0.1$. Como es de esperarse los valores de probabilidad son discretos y toman valores específicos para cada “número de no-conformes”. Nótese que el índice a asimetría es $\gamma = 0.27$ lo que indica una ligera asimetría positiva.

En la figura también se ha trazado la aproximación normal, y como se puede apreciar, el ajuste parece ser aceptable, especialmente en la región $\mu \pm 2\sigma$ (10 ± 6). Sin embargo, para gráficos de control p lo que interesa es el ajuste en los extremos de las colas, específicamente en los puntos determinados por los límites LCI y LCS.

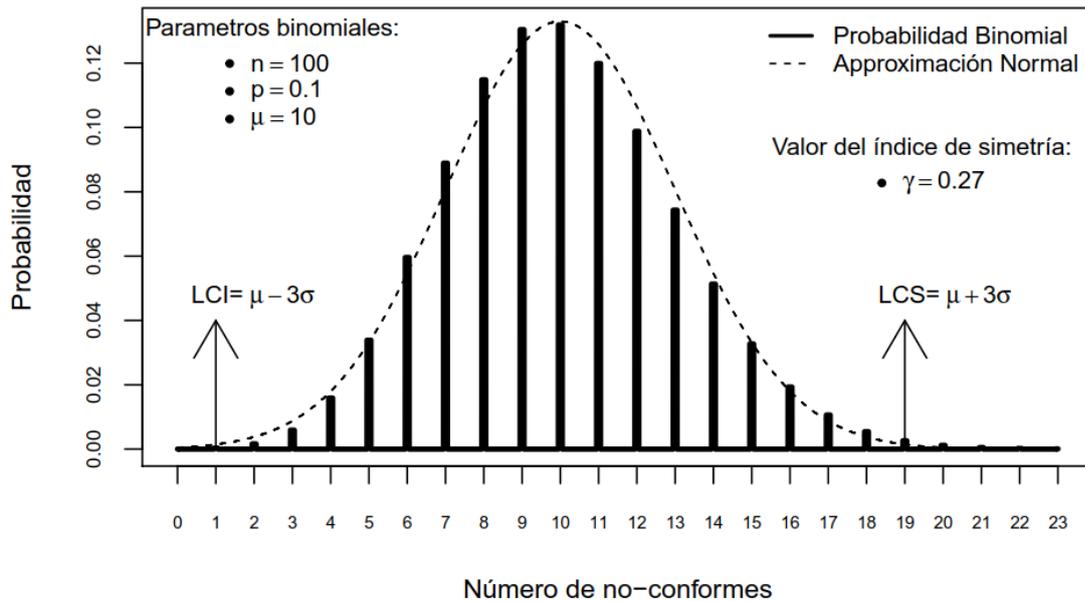


Figura 19. Ajuste de la aproximación normal a la binomial para $p = 0.1$ y $n = 100$

La Figura 20 muestra una ampliación de las colas de la Figura 19. Como se puede observar el ajuste de la aproximación normal es muy pobre en estos extremos de la distribución, esto a pesar de que la asimetría de la distribución binomial es pequeña y que el tamaño de muestra es relativamente alto. Nótese que en la cola inferior los valores discretos binomiales están por debajo de la curva normal, mientras que en la cola superior están por encima de ella, esto en buena parte es el motivo del por qué las colas alfa, para el GC p Shewhart, fluctúan por encima y por debajo de 0.00135.

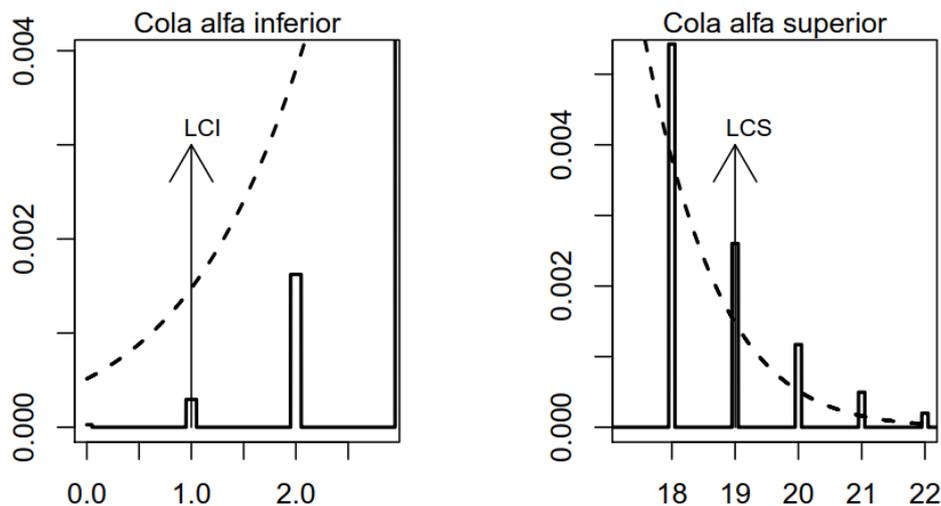


Figura 20. Aproximación normal a la binomial: ajuste en las colas para $n = 100$ y $p = 0.1$

Otro factor que influye en los valores de las colas alfa y que es muy importante tomar en cuenta, es la naturaleza discreta de la distribución binomial. Para explicar este factor se

incluyen la Figura 21 y la Figura 22, la primera muestra la distribución de probabilidad binomial acumulada para $n = 100$ y $p = 0.1$, mientras que la segunda muestra solo sus extremos (o colas). La forma escalonada de la curva es típica de una distribución discreta como la binomial.

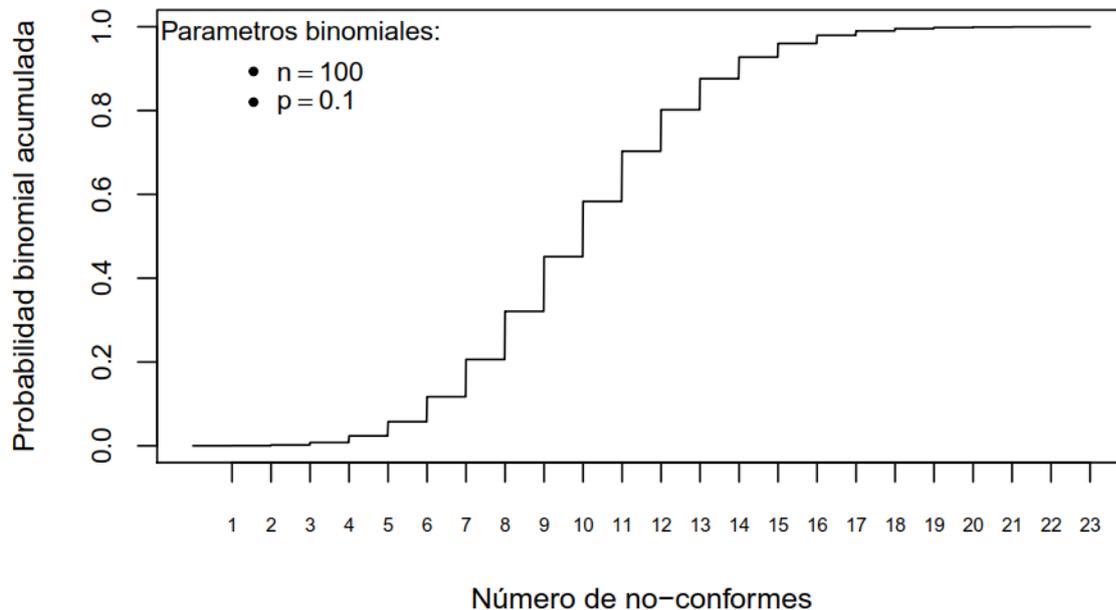


Figura 21. Probabilidad binomial acumulada para $n = 100$ y $p = 0.1$

En la Figura 22 se pueden apreciar los valores discretos de probabilidad que existen en las colas de la distribución. De ahí que la probabilidad de la cola alfa inferior que se obtiene con $nLCL=1$ sea 0.00032, que está muy por debajo de 0.00135, mientras que con $nUCL=19$ se obtiene una cola alfa superior igual a 0.00198, que está por encima de 0.00135. Nótese que en ninguna de las dos colas es posible obtener el valor nominal de 0.00135 y que el ratio de las colas alfa es $R_\alpha = 0.16$. Otro ejemplo del ajuste de la aproximación normal a la binomial se incluye en el Anexo 2. Estos ejemplos sirven para demostrar que los valores discretos que se dan en la distribución binomial varían y son específicos para cada combinación de p y n , y que debido a esto es poco probable que las dos colas alfa lleguen a obtener el valor nominal de 0.00135.

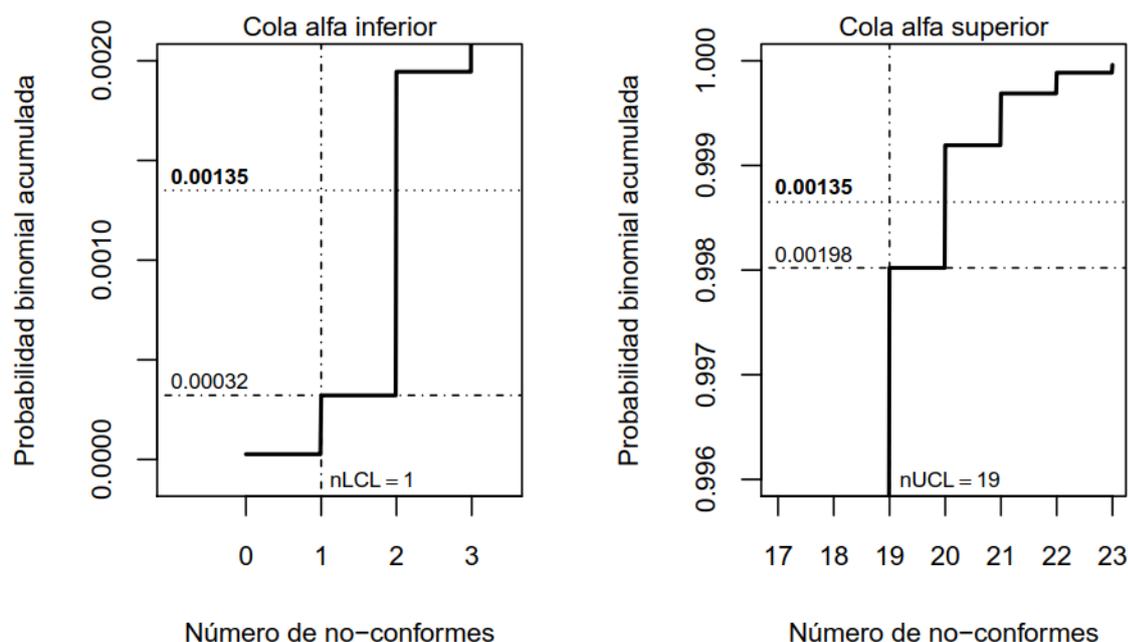


Figura 22. Colas de la probabilidad binomial acumulada para $n = 100$ y $p = 0.1$

4.1.7.1 El ratio de las colas alfa

Para tratar de cuantificar la disparidad que existe entre las colas alfa se creó el parámetro $R_\alpha = \alpha_L/\alpha_U$ que no es nada más que el ratio de las colas alfa. Un valor $R_\alpha \approx 1$ sería óptimo ya que indicaría que el valor de las dos colas es muy similar. Por otro lado, mientras más se aleje el R_α de 1 mayor será la disparidad entre las colas. En el ejemplo de la Figura 22 el $R_\alpha = \alpha_L/\alpha_U = 0.00032/0.00198 = 0.16$.

4.1.8 Gráfico de control p Shewhart : El ARL_1 y la curva ARL

En la sección 4.1.5.2 se presentó la ecuación (12), que sirve para calcular el ARL del proceso cuando este se encuentra fuera de control (o ARL_1). Utilizando esta ecuación se construyó la Tabla 5 que muestra los ARL_1 para cambios específicos en la fracción de no-conformes (p) del proceso, para el caso de $n = 100$ y $p = 0.1$.

Analizando los datos de la tabla, se puede ver que si p sufriese una modificación de -50%, es decir si cambiase de 0.1 a 0.05, el GC detectaría ese cambio y daría una alarma (en la forma de un punto fuera de los límites) en promedio tras 26 muestras. Mientras que si la modificación fuese del 100%, de 0.1 a 0.2, las alarmas se darían en promedio tras 2 muestras. Algo que llama mucho la atención son los valores ARL_1 para los cambios -20% y 20% debido a que son muy dispares, estos cambios, al ser de la misma magnitud, deberían detectarse con la misma rapidez, es decir su ARL_1 debería ser similar.

En general los resultados de la Tabla 5 parecen indicar que, para $n = 100$ y $p = 0.1$, un GC p Shewhart detectará, para una misma magnitud de cambio, más rápidamente incrementos de p (deterioros del proceso) que reducciones de p (mejoras del proceso).

Tabla 5: GC p Shewhart – ejemplo de valores ARL_1

Resultados para GC con parámetros $n = 100$ y $p = 0.1$			
Límites de control: LCL= 0.01 y LCS= 0.19			
Porcentaje de cambio en el proceso	p -modificada	$1 - \beta$	ARL_1
-50%	0.05	0.037081	26
-30%	0.07	0.006030	165
-20%	0.8	0.002434	410
20%	0.12	0.014747	67
30%	0.13	0.031952	31
50%	0.15	0.106545	9
100%	0.2	0.53983	2

Aunque la Tabla 5 es de mucha utilidad, una alternativa más eficaz para evaluar la capacidad de un GC para detectar cambios, es trazar los valores ARL_1 en una gráfica para así obtener la curva ARL, como se muestra en la Figura 23. Las curvas ARL son muy efectivas al momento de analizar la capacidad de detección de un GC, ya que ponen de manifiesto características que serían difíciles de apreciar en una tabla de datos como la Tabla 5. Acosta-Mejía (1999) fue uno de los primeros investigadores en utilizar las curvas ARL.

Para facilitar la interpretación de la curva se ha hecho que el eje- x sea igual a $(p_1/p) - 1$ donde $p_1 = p$ -modificado y $p =$ fracción de no-conformes del proceso. Cuando $(p_1/p) - 1 = 0$ el p no ha sufrido modificación alguna por lo que su valor correspondiente en el eje- y (ARL) da el ARL_0 del proceso. Valores $(p_1/p) - 1 < 0$ indican reducciones porcentuales de p , así por ejemplo $(p_1/p) - 1 = -0.2$ indica una disminución de p del -20%. Valores $(p_1/p) - 1 > 0$ indican incrementos porcentuales de p .

La Figura 23 muestra que el GC p Shewhart para $p = 0.1$ y $n = 100$ tiene una **curva ARL sesgada** hacia $(p_1/p) - 1 < 0$. La forma de la curva indica que este GC tardara más en detectar reducciones de p (mejoras del proceso) que incrementos (deterioros del proceso). Lo ideal sería que la curva fuese insesgada y que el ARL_0 estuviese en el punto máximo de la

curva, es decir que $ARL_0 = ARL_M$. Cabe destacar que una curva insesgada denota que un GC tiene la misma capacidad para detectar tanto mejoras como deterioros del proceso.

Cuando el sesgo es hacia $(p_1/p) - 1 < 0$, diremos que la curva tiene sesgo negativo, mientras que si el sesgo es hacia $(p_1/p) - 1 > 0$ diremos que la curva tiene sesgo positivo.

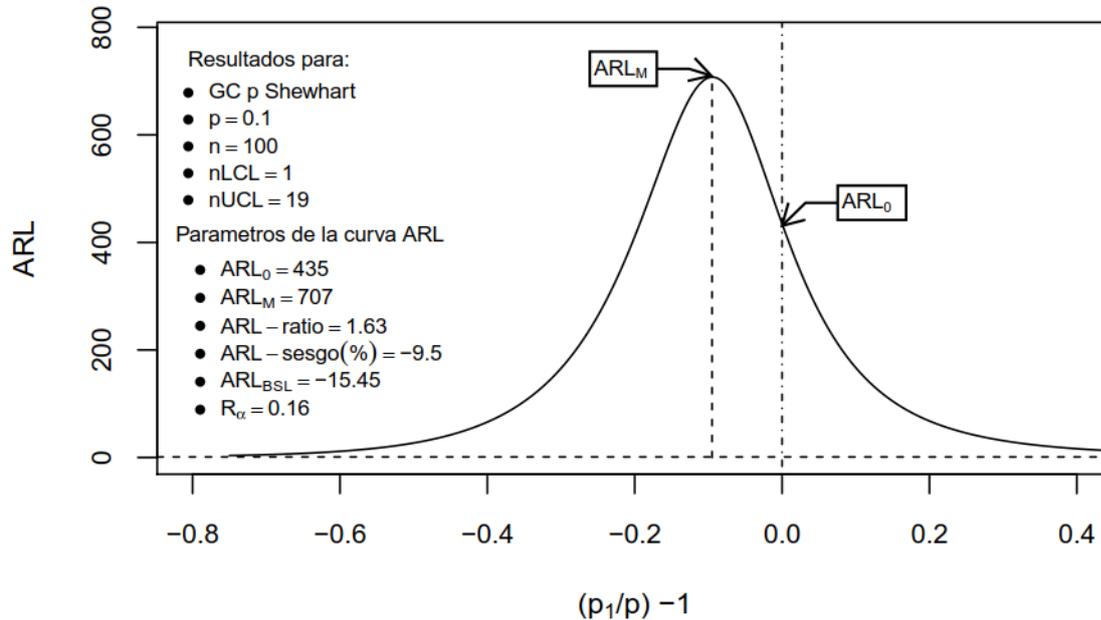


Figura 23. Curva ARL para GC p Shewhart con $n = 100$ y $p = 0.1$

La Figura 23 incluye varios parámetros que están agrupados bajo el título “*Parámetros de la curva ARL*” y que sirven para determinar las características de la curva, así como la severidad de su sesgo. En la Tabla 6 se define cada uno de esos parámetros,

4.1.8.1 Parámetros para determinar la severidad del sesgo en la curva ARL

En la Tabla 6 se definen todos los parámetros que se utilizan en esta tesis para determinar la severidad del sesgo en las curvas ARL. Sin embargo, hay que aclarar que los parámetros ARL_0 , ARL_M y $ARL\text{-sesgo}(\%)$ ya existían antes de comenzar el trabajo de investigación, ver (Morais, 2016), y que en un comienzo fueron estos los que se utilizaron para analizar el sesgo de la curva ARL. Los parámetros $ARL\text{-ratio}$, R_α y ARL_{BSL} fueron concebidos a lo largo de la investigación, por lo cual son novedades exclusivas de esta tesis.

Tabla 6: Definición de los parámetros de la curva ARL

- $ARL_0 = ARL$ del proceso cuando este está bajo control.
 - Algo importante de recordar es que valores muy altos o muy bajos del ARL_0 son indeseables y que en esta tesis se considera que el ARL_0 es aceptable si cae dentro del rango 250 a 450.

<ul style="list-style-type: none"> • ARL_M = Valor máximo de la curva.
<ul style="list-style-type: none"> • El ARL-ratio = ARL_M/ARL_0 <ul style="list-style-type: none"> – Idealmente el $(ARL_M/ARL_0) = 1$, caso contrario mientras más grande sea el ratio mayor será la severidad del sesgo.
<ul style="list-style-type: none"> • $ARL\text{-sesgo}(\%) = 100 * [(p_M/p) - 1]$ <p style="text-align: center;">$p_M = p$-modificado donde sucede el ARL_M.</p> <ul style="list-style-type: none"> – El ARL-sesgo(%), también llamado sesgo ARL relativo, es la diferencia proporcional entre la ubicación del ARL_M y el ARL_0 en el eje-x. – Idealmente el ARL-sesgo(%) debería ser igual a cero. Mientras más se aleje de cero mayor será la severidad del sesgo. – Un ARL-sesgo(%) negativo indica que la curva tiene sesgo negativo.
<ul style="list-style-type: none"> • ARL_{BSL} = ARL bias severity level (nivel de severidad del sesgo) <ul style="list-style-type: none"> – ARL_{BSL} sirve para cuantificar la severidad de sesgo en la curva ARL. Se calcula mediante la ecuación (26), ver sección 5.4.
<ul style="list-style-type: none"> • $R_\alpha = \alpha_I/\alpha_S$ <ul style="list-style-type: none"> – R_α es el ratio de las colas alfa.

4.1.8.2 Gráfico de control p Shewhart: variación de la severidad del sesgo en las curvas ARL

La Figura 24 muestra a modo de ejemplo, las curvas ARL para $p = 0.05$ con $n = 220$ y 600 . Las dos curvas tienen sesgo negativo, aunque se aprecia claramente que el sesgo de la curva $n = 220$ es mucho más severo que la de $n = 600$. Más ejemplos de curvas ARL se incluyen en el Anexo 3. Todos esos ejemplos sirven para ilustrar algo que llamó mucho la atención durante los análisis iniciales de curvas ARL, siendo esto que, para $p \ll 0.5$, todas las curvas parecían tener sesgo negativo cuya severidad variaba según el valor de los parámetros binomiales p y n .

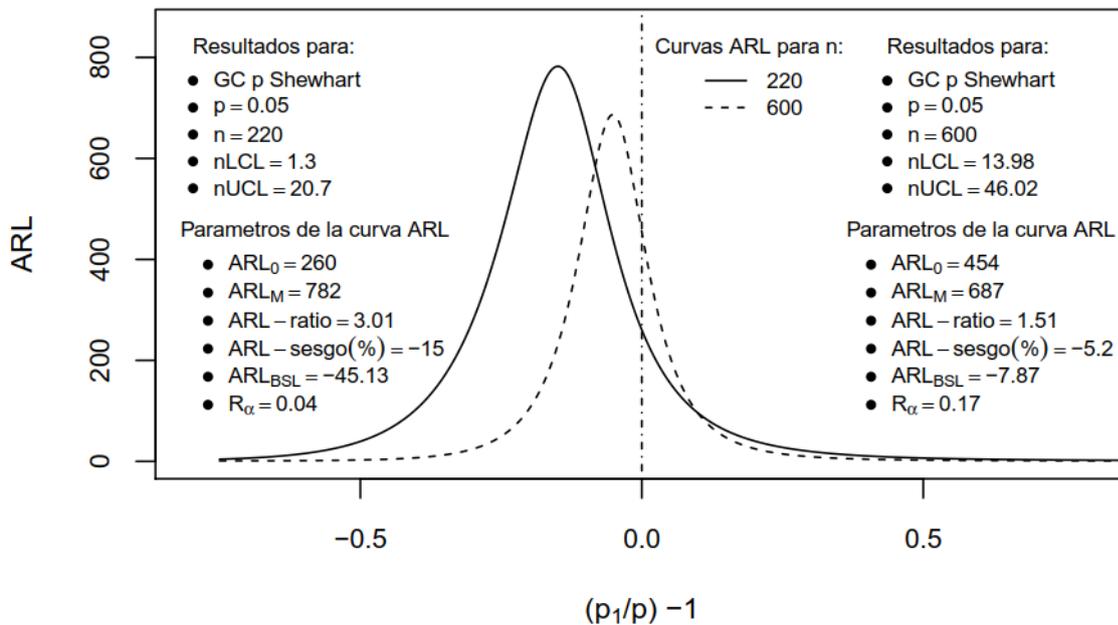


Figura 24. Curvas ARL para $p = 0.05$ y $n = 220$ y 600

Algo importante de destacar en este punto, son los valores del R_α en las curvas ARL presentadas anteriormente. Nótese que mientras más alejado es el valor R_α de 1, mayor es el valor del ARL-sesgo(%) y que esto parece ir acompañado de un incremento en la severidad del sesgo. En la fase exploratoria de este trabajo de investigación, estas observaciones hechas en la curva ARL, parecían indicar que había una relación directamente proporcional entre la severidad del sesgo y el R_α . Para confirmar si esto era verdad, y con el fin de determinar la relación que existía entre el ARL-sesgo(%) y R_α , se realizó un estudio cuyos objetivos fueron:

- 1: Caracterizar el comportamiento del ARL-sesgo(%) en función de los parámetros binomiales p y n .
- 2: Determinar la relación entre el ARL-sesgo(%) y el R_α en función de p y n .

Los resultados de este estudio se resumen en la sección 5.2.

4.1.9 Sinopsis de la sección 4.1

La sección 4.1 resume exclusivamente el marco teórico que sirvió de base para determinar el efecto que tiene, en la capacidad de monitorización de GCs tipo Shewhart, el impreciso ajuste de la aproximación normal a la distribución binomial. Todo el trabajo de investigación relevante y posterior a lo presentado en esta sección, se incluye en la sección 5.2.

4.2 Gráficos de control por atributos basados en la distribución de Poisson

Al final de la sección 3.2.2.2 se mencionó que la pregunta que había dado inicio a la investigación resumida en esta tesis era: *¿Cómo afecta la imprecisión de la aproximación normal a la capacidad de monitorización de los GCs por atributos tipo Shewhart?* En esta sección se resume el marco teórico que sirvió de base para responder esa pregunta, para el caso específico de GCs basados en la distribución de Poisson.

4.2.1 Introducción al uso de la distribución de Poisson en gráficos de control por atributos

En muchos casos es necesario, o más conveniente, monitorizar un proceso en base al número de defectos (también llamados no-conformidades) que tienen los bienes producidos, que mediante el número de bienes no-conformes. Para explicar el tipo de procesos en los que es preferible utilizar “defectos” y el por qué se utiliza la distribución de Poisson para estos casos, se dan un par de ejemplos.

Supongamos que una fábrica produce ordenadores y que antes de que estos salgan a la venta, se realiza un muestreo cuya finalidad es identificar el número de “defectos” que tienen a partir de atributos de calidad y/o especificaciones pre establecidas. Si la cantidad de defectos son pocos y leves, entonces la calidad del ordenador podría ser considerada como aceptable, lo que llevaría a que sea valorado como un producto apto para la venta. En cambio si la cantidad de defectos es elevada, el ordenador podría ser considerado de baja calidad y por ende valorado como un producto no apto para la venta.

En el ejemplo expuesto en el párrafo anterior, no se clasifica al computador simplemente como “apto” o “no apto” (también podría ser conforme o no-conforme) debido a que no cumple con un solo atributo de calidad. Más bien su clasificación depende de la cantidad de defectos (o no conformidades) que tenga. Es importante aclarar que los defectos podrían darse aleatoriamente en cualquier parte del ordenador, esto implica que el número de oportunidades (o ubicaciones) donde podrían aparecer dichos defectos es muy grande.

Otro ejemplo es el número de defectos por x -metros cuadrados de tela, en este caso se considera que la tela es apta para la venta solo cuando la cantidad de defectos por x -metros cuadrado no sobrepasa un cierto número. En este caso también se da que el número de oportunidades (o ubicaciones) donde podrían aparecer los defectos es muy grande.

En los ejemplos anteriores el número de defectos (sea en uno o en un grupo de computadores o en x -metros cuadrados de tela), puede ser considerado como una variable aleatoria que sigue una distribución de probabilidad del tipo Poisson, cuya probabilidad de ocurrencia está dada por la ecuación (14), siempre y cuando la ocurrencia de defectos sea pequeña en comparación con el número de oportunidades (o ubicaciones) donde podrían aparecer dichos defectos. Si esto se cumple, entonces procesos que producen bienes como los expuestos en los ejemplos anteriores, pueden ser monitorizados mediante gráficos de control basados en esa distribución.

Por ejemplo, en el caso de la tela, podría ser que en promedio se dé un solo defecto por cada diez metros cuadrados de tela, por lo que en 100 metros cuadrados se darían, en promedio, solo 10. Cien metros cuadrados representan una área extensa en donde se podrían presentar fallas, por lo que tener solo diez es muy bajo, por eso es que en este caso se podría aplicar la distribución de Poisson. Montgomery (2009, pág. 308) y Walpole et al (2007, pág. 161) dan varios ejemplos de este tipo de procesos.

4.2.2 Propiedades de la distribución de Poisson - $Poi(c)$

Si el número de defectos en una muestra se modela como una variable aleatoria X que sigue una distribución de Poisson $X \sim Poi(c)$, entonces mediante la ecuación (14) se podría determinar la probabilidad de que ocurran un determinado número de defectos. La ecuación (14) es la función de probabilidad de Poisson.

$$P(x, c) = \frac{e^{-c} c^x}{x!} \quad x = 0, 1, 2, 3 \dots \dots \dots \quad (14)$$

Donde:

- x = número de defectos
- c = promedio de defectos

Para la distribución de Poisson el valor promedio y la varianza coinciden, como se muestra en ecuación (15).

$$E[X] = Var[X] = c \quad (15)$$

c = promedio de defectos

Por otro lado el coeficiente de simetría para la distribución Poisson está dado por (16). En este caso, al considerar la raíz cuadrada con signo positivo, la asimetría es siempre positiva.

$$\gamma = \frac{1}{\sqrt{c}} \quad (16)$$

Dónde:

$\gamma = 0$, indica una distribución totalmente simétrica

$\gamma > 0$, indica una distribución asimétrica positiva

4.2.3 Aproximación normal a la distribución de Poisson

Es conocido que cuando n tiende a infinito ($n \rightarrow \infty$) y p tiende a cero ($p \rightarrow 0$) la distribución binomial converge hacia la distribución de Poisson. A continuación se comprueba esa convergencia.

Las funciones generadoras de momentos para la distribución binomial $X \sim B(n, p)$ y la distribución de Poisson $Y \sim Poi(c)$ están dadas por (p.1) y (p.2) respectivamente.

$M_x(t) = (q + pe^t)^n$ *ver (b.4) en la sección 4.1.2.2

$$M_x(t) = [pe^t + (1 - p)]^n \quad (p.1)$$

$$M_Y(t) = e^{c(e^t-1)} \text{ donde } c = np \text{ y } p \rightarrow 0 \quad (p.2)$$

(p.1) puede ser re escrita como (p.3), y ya que $c = np$, también como (p.4). Si asumimos que en (p.4) n es muy grande, entonces es posible aplicar el lema mostrado en (p.5) a la expresión (p.4) lo que da como resultado (p.6) que muestra la convergencia de la distribución binomial a la Poisson.

$$M_x(t) = [1 + \frac{1}{n}np(e^t - 1)]^n \quad (p.3)$$

$$M_x(t) = [1 + \frac{1}{n}c(e^t - 1)]^n = [1 + \frac{c(e^t-1)}{n}]^n \quad (p.4)$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{a}{n}\right)^n = e^a \quad (p.5)$$

$$\lim_{n \rightarrow \infty} M_x(t) = e^{c(e^t-1)} = M_Y(t) \quad (p.6)$$

Donde $c = np$ tendría un valor finito debido a que $p \rightarrow 0$

De esta comprobación se puede concluir que la distribución de Poisson en un caso especial de la distribución binomial para cuando ($n \rightarrow \infty$) y ($p \rightarrow 0$). A partir de este antecedente entonces también es lógico asumir que la distribución de Poisson, al igual que la binomial, puede ser aproximada por la distribución normal. De ahí que la ecuación (17) puede ser

utilizada en vez de la función de probabilidad Poisson, dada en la ecuación (14), para estimar la probabilidad de que ocurran un determinado número de defectos por muestra. Información más detallada acerca de la convergencia y la aproximación normal a la distribución de Poisson la dan Simons & Johnson (1971) y Rich (2009) respectivamente.

$$P(x, c) \approx \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (17)$$

Dónde:

- $x = \text{variable aleatoria}$
- $\mu = c$
- $\sigma = \sqrt{c}$

4.2.3.1 Ajuste de la aproximación normal a la distribución de Poisson

En esta sección se analiza el ajuste de la aproximación normal a la distribución de Poisson. Una regla empírica que se aplica para determinar la condición que debe existir para obtener un buen ajuste de la aproximación normal a la distribución Poisson, dice que esto se da cuando $c \geq 10$, ver (Mitra, 2008, pág. 212).

La Figura 25 muestra la densidad de probabilidad para $c = 10$ y también su aproximación normal, como se puede apreciar el ajuste de la normal es aceptable entre el rango $\mu \pm 2\sigma$ (10 ± 6.32). Sin embargo, y como se puede observar en la figura Figura 26, el ajuste en las colas de la distribución es muy impreciso. Nótese que el índice de simetría indica que la distribución tiene una ligera asimetría positiva.

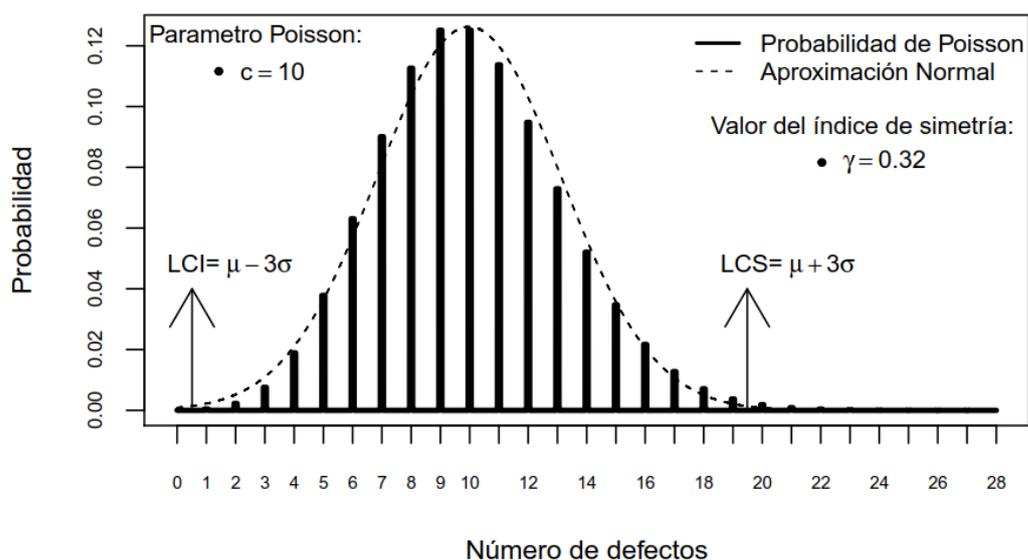


Figura 25. Probabilidad de Poisson para $c = 10$ y su aproximación normal

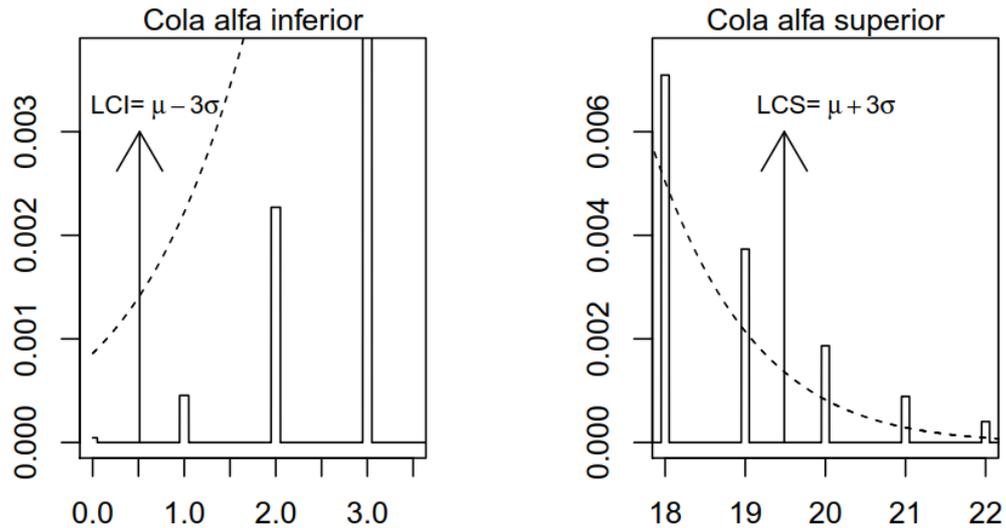


Figura 26. Colas de la probabilidad de Poisson para $c=10$ y su aproximación normal

La Figura 27 muestra la densidad de probabilidad para $c = 20$ y su aproximación normal. El índice de simetría es menor que el obtenido con $c = 10$, lo que indica que con $c = 20$ la distribución de Poisson es un poco más cercana a normal. Nótese que el ajuste de la distribución normal es aceptable entre el rango $\mu \pm 2\sigma$ (20 ± 8.9), pero que, como se muestra en la Figura 28, el ajuste en las colas de la distribución sigue siendo muy impreciso.

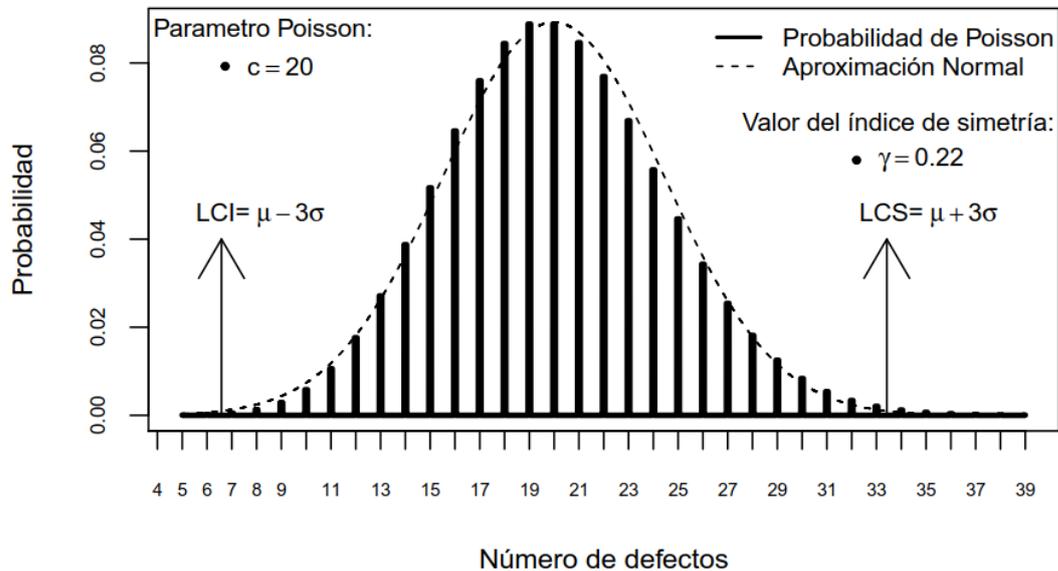


Figura 27. Probabilidad de Poisson para $c = 20$ y su aproximación normal

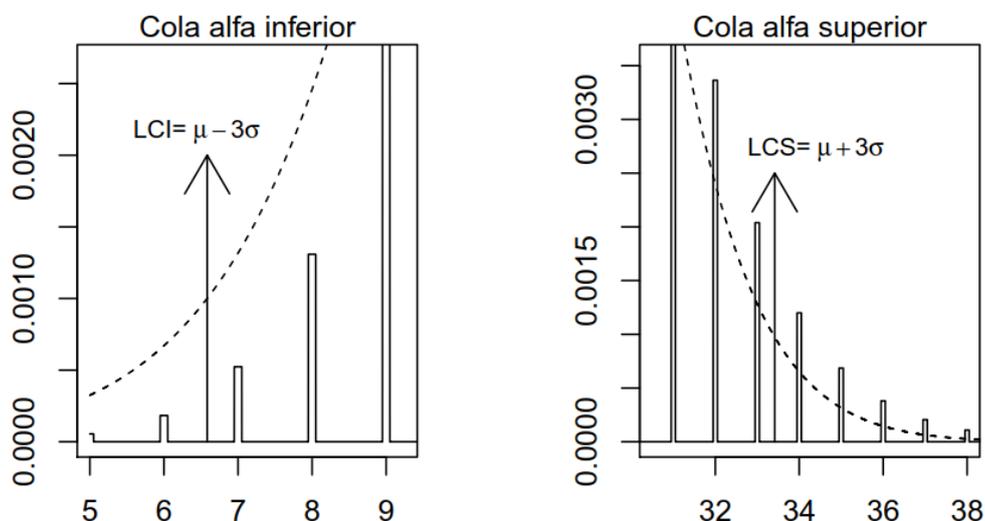


Figura 28. Colas de probabilidad de Poisson para $c = 20$ y su aproximación normal

Cabe resaltar que los puntos LCI y LCS ($\mu \pm 3\sigma$) indican el sitio bajo la curva normal donde el área de probabilidad es 0.00135. Nótese que en los ejemplos ningún valor discreto de probabilidad Poisson coincide con estos puntos, además es fácil darse cuenta que los valores discretos están: i) por debajo de 0.00135 para la cola inferior y ii) por encima de 0.00135 para la cola superior.

La imprecisión del ajuste de la aproximación normal a las colas de la distribución de Poisson, es similar a lo que pasa con el ajuste de la aproximación normal a la distribución binomial, ver sección 4.1.7.

4.2.4 Principios de gráficos de control c y u Shewhart

Para construir los gráficos de control (GC) c y u del tipo Shewhart, que es el tema de estudio de esta sección, se asume que la distribución de Poisson es aproximadamente normal, generalmente esto se lo comprueba con reglas empíricas como $c \geq 10$. Si esta regla se cumple entonces se asume que la desviación estándar de la $v.a.^6(x)$ [x = número de defectos en una muestra] es $\sigma_x = \sqrt{c}$, o también que la desviación estándar de la $v.a(x/n)$ [x/n = promedio de defectos por muestra] es $\sigma_{(x/n)} = \sqrt{u/n}$. Los límites de control se los calcula mediante las ecuaciones dadas en la Tabla 7. Montgomery (2009, pág. 309) detalla los principios básicos de estos GCs.

⁶ $v.a$ = variable aleatoria

Algo que hay que clarificar antes de continuar, es el concepto del tamaño de muestra (n) para el caso de control por defectos (GCs c y u). En este caso el tamaño de muestra n está relacionado con lo que se conoce como “unidad de inspección (UDI)”. Una UDI puede ser por ejemplo 10 m^2 de tela, en este caso $n = 1 = 1 * \text{UDI} = 10 \text{ m}^2$ de tela, si $n = 2.5$ entonces $2.5 * \text{UDI} = 25 \text{ m}^2$ de tela, o si $n = 0.5 = 5 \text{ m}^2$ de tela. Como se puede apreciar una UDI equivale a $n = 1$ y n puede tomar valores menores o mayores que uno dependiendo del caso.

Tabla 7: Ecuaciones para calcular los límites de control para los gráficos c y u tipo Shewhart

<i>Gráfico de control c</i>	<i>Gráfico de control u</i>
<p style="text-align: center;">Linea central = c</p> <p style="text-align: center;">$LCS = c + (K_{\alpha/2}) \sigma_x$</p> <p style="text-align: center;">$LCI = c - (K_{\alpha/2}) \sigma_x$</p> <p>Dónde:</p> <ul style="list-style-type: none"> • c = Media del número de defectos en una UDI. • $\sigma_x = \sqrt{c}$ • $(K_{\alpha/2}) = 3$ • LCS = Limite de control superior • LCI = Limite de control inferior 	<p style="text-align: center;"><i>Linea central = u</i></p> <p style="text-align: center;">$LCS = u + (K_{\alpha/2}) \sigma_{(x/n)}$</p> <p style="text-align: center;">$LCI = u - (K_{\alpha/2}) \sigma_{(x/n)}$</p> <p>Dónde:</p> <ul style="list-style-type: none"> • u = Promedio de defectos por UDI. <p>Para determinar u:</p> $u = \frac{\sum_{i=1}^m y_i}{\sum_{i=1}^m n_i}$ <p>y_i = número de defectos en la i-ésima muestra</p> <p>$n_{(i=1,2,3...m)}$ = Una cantidad predeterminada de m muestras.</p> <ul style="list-style-type: none"> • $\sigma_{(x/n)} = \sqrt{\frac{u}{n}}$ • $(K_{\alpha/2}) = 3$

Como se puede apreciar en la Tabla 7, los límites de control se los ubica a una distancia de tres desviaciones estándar ($\pm 3\sigma$) de la media. También hay que recalcar que, al igual que para los gráficos p y np , cada vez que se traza un punto en el GC se esta en efecto realizando una prueba de hipótesis de dos colas (ver Tabla 4). Por lo que, al fijar los límites a $\pm 3\sigma$, la probabilidad total de cometer un error de tipo I (α) es de $\alpha = 0.0027$ o $\alpha/2 = 0.00135$ por cola. Cabe recordar que la definición del término colas alfa utilizado en esta tesis, se encuentra en la sección 4.1.4.2 y que este hace referencia a las probabilidades de que, en un

gráfico de control, un punto caiga fuera de uno de los límites (LCI o LCS) cuando el proceso está bajo control.

Según Ryan (2011, pág. 211) cuando se utiliza un tamaño de muestra (n) constante se podría utilizar tanto el gráfico c como el u , mientras que si n es variable se debería usar el gráfico u . De este punto en adelante, para no redundar y ser conciso, se utilizará sólo el GC u asumiendo n constante para resumir la investigación desarrollada para los GCs basados en la distribución de Poisson. Aunque los resultados, como se explica en la sección 5.6.3.6, también aplican al caso de n variable.

4.2.4.1 ARL para el gráfico de control u cuando el proceso está bajo control (ARL_0)

Basado en la teoría desarrollada en la sección 4.1.5 sabemos que, cuando un proceso está bajo control el ω de ecuación (7) es igual a la probabilidad de cometer un error de Tipo I (α), de esto se obtiene ecuación (18) que es igual a la ecuación (8) utilizada para el GC p .

$$ARL_0 = \frac{1}{\omega} = \frac{1}{\alpha} \quad (18)$$

El ARL_0 también puede ser calculado mediante la ecuación (19), esto debido a que la probabilidad de error tipo I (α) es igual a la suma de las probabilidades de las colas alfa, ver 4.1.4.2.

Cabe recordar que el ARL_0 representa la media de puntos trazados dentro de los límites de control, antes de que uno caiga fuera de ellos. En este caso, y ya que el proceso está bajo control, el punto que cae fuera de los límites se considera como una falsa alarma.

$$ARL_0 = \frac{1}{\alpha} = \frac{1}{\alpha_I + \alpha_S} \quad (19)$$

Dónde:

$\alpha_S =$ Cola alfa superior (Probabilidad de que un punto caiga fuera del LCS cuando el proceso está bajo control)

$\alpha_I =$ Cola alfa inferior (Probabilidad de que un punto caiga fuera del LCI cuando el proceso está bajo control)

Para el caso del gráfico de control u , las áreas de las colas se las calcula utilizando la probabilidad de Poisson acumulada, como se muestra en las ecuaciones (20) y (21).

$$\alpha_S = P\{x > nLCS |_{c=nu}\} = 1 - \sum_{x=0}^{x=nLCS} \frac{e^{-c} c^x}{x!} \quad x = 0,1,2 \dots nLCS \quad (20)$$

$$\alpha_I = P\{x < nLCL|_{c=nu}\} = \sum_{x=0}^{x=nLCL} \frac{e^{-c} c^x}{x!} \quad x = 0,1,2 \dots nLCS \quad (21)$$

4.2.4.2 ARL para el gráfico de control u cuando el proceso está fuera de control (ARL_1)

Cuando el proceso ha sufrido un cambio, es decir el valor de u se ha modificado debido a un incremento o disminución en el promedio de defectos por UDI, se tiene que $\omega = 1 - \beta$, donde β es la probabilidad de cometer un error de Tipo II (ver sección 4.1.5). Por lo que el ARL para este caso, y que en esta tesis se denota como ARL_1 , está dado por la ecuación (22).

Cabe recordar que ARL_1 representa el promedio de puntos trazados dentro de los límites de control antes de uno caiga fuera de ellos, cuando el proceso está fuera de control.

$$ARL_1 = \frac{1}{1 - \beta} \quad (22)$$

Dónde:

$$\beta = Pr\{x \leq nUCL|_{c=nu_1}\} - Pr\{x \leq nLCL|_{c=nu_1}\}$$

$u_1 = u$ -modificado

u_1 podría ser: $u_1 < u$ o $u_1 > u$

$Pr =$ Probabilidad de Poisson acumulada

4.2.5 El ARL_0 en gráficos de control u Shewhart

En la Tabla 7 se presentó la ecuación para calcular los límites de control para el gráfico de control u Shewhart. Cabe recordar que esa ecuación ha sido derivada con base en la presunción de que la aproximación normal a la distribución de Poisson es adecuada. La regla empírica $c \geq 10$ (o $nu \geq 10$) es comúnmente utilizada para verificar la idoneidad de esa aproximación.

Si la aproximación es adecuada entonces lo común es asumir que cada una de las colas alfa tiene un valor, sino igual, al menos muy cercano a 0.00135, ver ecuación (4). Por lo que mediante la ecuación (19) se obtendría que el $ARL_0 = 1/(0.00135 + 0.00135) = 1/(0.0027) \approx 370$. Lo que significaría que cuando el proceso está bajo control, en promedio, 370 puntos caerían dentro de los límites de control antes de que uno cayese fuera de ellos. Recordemos que los puntos que caen fuera de los límites cuando el proceso está bajo control, son considerados falsas alarmas.

Sin embargo, y como se demuestra más adelante, la presunción de que las colas alfa del GC u Shewhart son iguales a 0.00135 es errónea y por lo tanto asumir que $ARL_0 \approx 370$ también lo es. Esto se debe principalmente a que la distribución de Poisson es altamente asimétrica, lo que causa que el ajuste de la aproximación normal sea muy inexacto en las colas de esa distribución, tal como se demostró en la sección 4.2.3.1.

En las siguientes secciones se analiza el comportamiento real de las colas alfa y también su efecto en el ARL_0 del GC u .

4.2.5.1 Comportamiento real de las colas alfa para el gráfico de control u Shewhart

Vamos a suponer que un proceso tiene un $u = 2$ y un tamaño de muestra $n = 10$. Con estos datos se obtiene que los límites de control para el GC u son $LCI = 0.658$ y $UCL = 3.34$. En este caso se obtiene que $c = nu = (10 \cdot 2) = 20$, que está muy por encima de la regla empírica $nu \geq 10$, en este caso la práctica habitual es simplemente asumir que la aproximación normal a la de Poisson es adecuada y que en consecuencia el ARL_0 es 370. Sin embargo utilizando las ecuaciones (20) y (21), mediante las cuales se puede calcular los verdaderos valores de las colas alfa, se obtiene que $\alpha_L = 0.00025$ y $\alpha_S = 0.00268$ por lo que el verdadero ARL_0 es 340. Como se puede apreciar los verdaderos valores de las colas alfa no son iguales a 0.00135, además llama mucho la atención la gran diferencia que existe entre las cola alfa y el valor nominal 0.00135.

El hecho de que las colas alfa no fuesen iguales a 0.00135, como se demuestra en el ejemplo del párrafo anterior, motivó la realización de un estudio que tuvo con fin caracterizar el comportamiento de las colas alfa del GC u Shewhart. Ejemplos de los resultados obtenidos se exponen en la Figura 29 y en las figuras del Anexo 7, estas figuras muestran los verdaderos valores de las colas alfa para $u = 1, 2$ y 3 dentro de rangos de n específicos para cada u . Los rangos de n fueron determinados mediante la ecuación (23) la cual es comúnmente utilizada para determinar el tamaño de muestra para el GC u , ver (Duncan, 2000, pág. 460).

Para establecer el valor mínimo y máximo de n , se aplicaron los siguientes criterios: i) que con todas las combinaciones de n y u se debían obtener límites de control inferior (LCI) y superior (LCS) y ii) que el tamaño de muestra debía ser lo suficientemente grande. Para cumplir estos criterios se utilizó en (23) $d = 0.97u$ para el valor n mínimo y $d = u/2$ para el n máximo. En los gráficos el rango de n se muestra como: n rango = n mínimo (0.1) n máximo, por ejemplo en la Figura 29 se muestra n rango = 10(0.1)36, indicando que los cálculos fueron realizados desde el tamaño de muestra 10 hasta el 36 en incrementos de 0.1.

$$n = 9u/d^2 \quad (23)$$

Al igual que lo dicho para el GC p Shewhart en la sección 4.1.6.1, para este caso también hay que aclarar la diferencia que existe entre los conceptos “rango de tamaños de muestra” y “tamaño de muestra variable”. El primero hace referencia a una secuencia de n que tiene un mínimo y máximo y cuyo valor incrementa acorde a una magnitud específica pre-establecida, por ejemplo $10(0.1)36$, mientras que el segundo se refiere a un número no específico de muestras n_i ($i = 1, 2, 3 \dots$) cuyos tamaños pueden variar arbitrariamente, por ejemplo $n_1 = 2$, $n_2 = 1$, $n_3 = 1.5$ etc. El concepto de “tamaño de muestra variable” es utilizado para el caso de monitorización de procesos mediante muestras no constantes, mientras que “rango de tamaños de muestra” es utilizado en esta tesis para analizar una serie, o grupo, de gráficos individuales ordenados en una secuencia continua. Por ejemplo para el caso $u = 1$ y n rango = $10(0.1)36$, se tienen un grupo de 260 gráficos, es decir un gráfico para cada combinación de u y n .

De los resultados expuestos en la Figura 29 y en el Anexo 7, se puede concluir que las colas alfa del GC u Shewhart no tienen un valor constante, sino que estas más bien oscilan por encima (cola superior) y por debajo (cola inferior) de 0.00135, y que además la cola alfa inferior se acerca a cero a medida que el tamaño de muestra se reduce. Algo que también es posible deducir es que los valores de las colas alfa están en función de los parámetros u y n .

Nótese que si se aplicara la regla empírica $nu \geq 10$ a cualquiera de las combinaciones n y u de los gráficos de resultados, nos llevaría a la errónea conclusión de que la aproximación normal a la distribución de Poisson es adecuada y que en consecuencia las colas alfa tendrían un valor, sino igual, al menos muy cercano a 0.00135. Por esta razón se concluye que esta regla empírica no es adecuada para establecer la idoneidad de las colas alfa del GC u Shewhart.

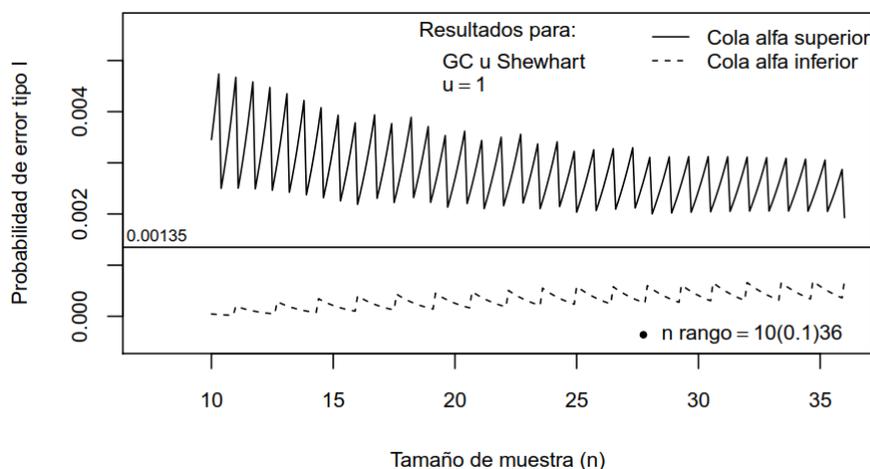


Figura 29. GC u Shewhart : Comportamiento real de la colas alfa para $u = 1$

4.2.5.2 Comportamiento real del ARL_0 en el gráfico de control u Shewhart

La Figura 30 muestra los ARL_0 correspondientes a las colas alfa de la Figura 29 y que fueron obtenidos mediante la ecuación (19). Como se puede apreciar el ARL_0 en vez de tener un valor constante, mas bien oscila alrededor de 370. Además, en la mayoría de ocasiones, su valor esta por debajo de 370. Resultados del ARL_0 para $u = 2$ y 3 se incluyen en el Anexo 7.

Los resultados demuestran que la fluctuación de las colas alfa afecta directamente al ARL_0 y que su valor real esta muy alejado de 370. Además, también es lógico deducir que el ARL_0 esta en función de u y n . La implicación de esto es que cuando se utiliza el GC u Shewhart existe una alta probabilidad de que el ARL_0 no sea 370. En este caso lo adecuado sería computar el ARL_0 y determinar si su valor es aceptable.

Ahora bien, para poder establecer si un ARL_0 es o no aceptable, hay que tomar en cuenta que tanto valores muy altos como muy bajos del ARL_0 son indeseables para una correcta monitorización del proceso. Un ARL_0 muy alto podría afectar la rapidez del gráfico para detectar cambios en el proceso, mientras que un ARL_0 muy bajo daría lugar a que se presenten falsas alarmas con demasiada frecuencia. En esta tesis, y basado en lo expuesto en la sección 5.1.3, se considera que el ARL_0 para los GCs u es aceptable si cae dentro del rango 250 a 450.

Nótese en la Figura 30 que el 94% de los gráficos u Shewhart contruidos con $u = 1$ y $n = 10(0.1)36$ tendrían un ARL_0 aceptable.

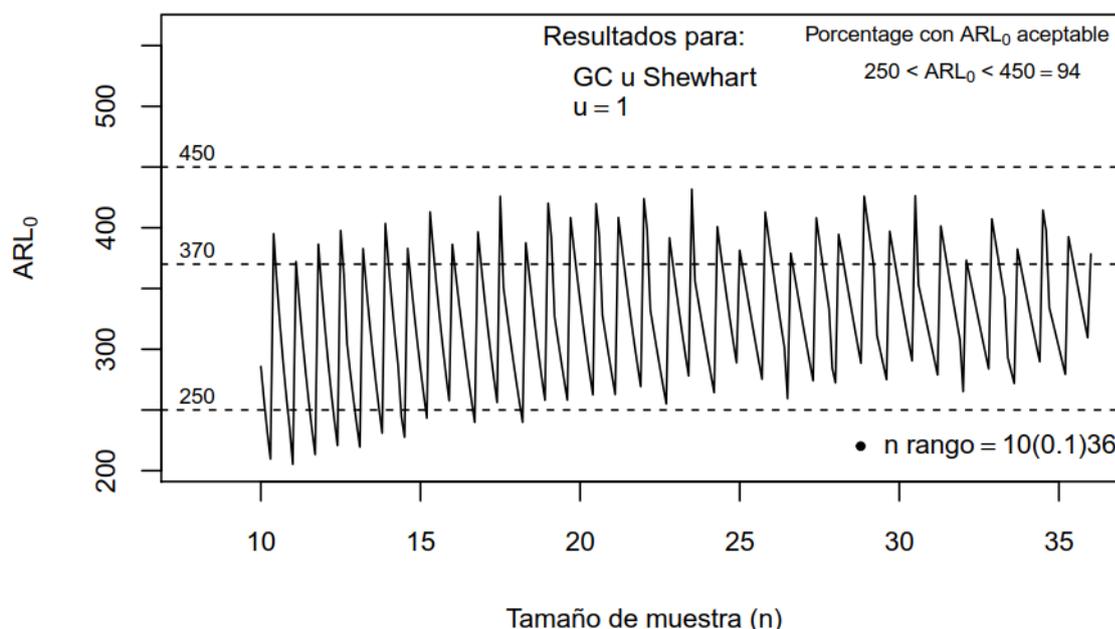


Figura 30. GC u Shewhart: Oscilación del ARL_0

4.2.6 La curva ARL para el gráfico de control u Shewhart

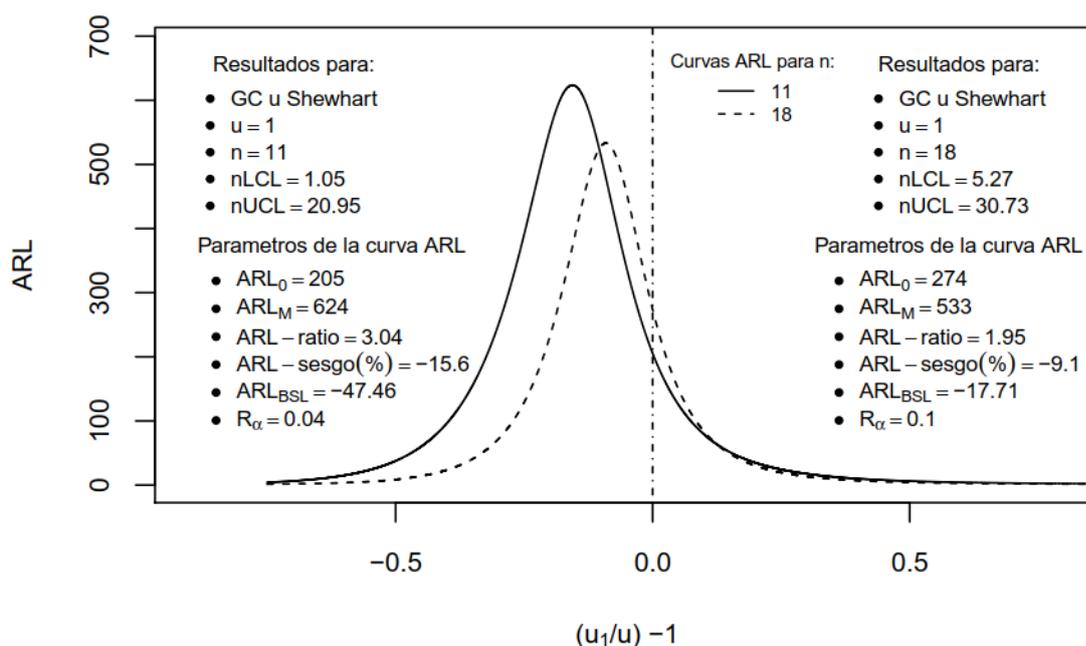


Figura 31. GC u Shewhart: Curvas ARL para $u = 1$ con $n = 11$ y $n = 18$

Igual que para el GC p , la capacidad del GC u para detectar cambios en el proceso también puede ser fácilmente analizada mediante la curva ARL. Estas curvas se las obtiene mediante la ecuación (22). La Figura 31 muestra un par de ejemplos de curvas ARL para gráficos u Shewhart para $u = 1$ con $n = 11$ y $n = 18$.

Para facilitar la interpretación de la curvas ARL se ha hecho que el eje- x sea igual a $(u_1/u) - 1$ donde $u_1 = u$ -modificado y $u =$ promedio de defectos por UDI⁷. Cuando $(u_1/u) - 1 = 0$ el u no ha sufrido modificación alguna por lo que su valor correspondiente en el eje- y (ARL) da el ARL_0 del proceso. Valores $(u_1/u) - 1 < 0$ indican reducciones porcentuales de u , así por ejemplo $(u_1/u) - 1 = -0.2$ indica una disminución de u del -20%. Valores $(u_1/u) - 1 > 0$ indican incrementos porcentuales de u .

La Figura 31 muestra que los gráficos u Shewhart para $u = 1$ y $n = 11$ y 18 tienen **curvas ARL sesgadas** hacia $(u_1/u) - 1 < 0$. La forma de la curva denota que estos gráficos tardaran más en detectar reducciones de u (mejoras del proceso) que incrementos de u (deterioros de proceso). Lo ideal sería que la curva fuese insesgada y que el ARL_0 estuviese en el punto máximo de la curva. Cuando el sesgo es hacia $(u_1/u) - 1 < 0$, diremos que la curva tiene

⁷ UDI= Unidad de inspección

sesgo negativo, mientras que si el sesgo es hacia $(u_1/u) - 1 > 0$ diremos que la curva tiene sesgo positivo.

La Figura 31 incluye varios parámetros agrupados bajo el título “*Parámetros de la curva ARL*”. Estos parámetros son los mismos que los utilizados para las curvas ARL de los GCs p . Las definiciones de cada parámetro se encuentran en la Tabla 6 (ver sección 4.1.8).

Otros ejemplos de curvas ARL para el GC u Shewhart se encuentran en el Anexo 8 en donde se puede observar que todas las curvas tienen sesgo negativo.

4.2.6.1 Gráfico de control u Shewhart: Severidad del sesgo en las curvas ARL

En los análisis preliminares de curvas ARL provenientes del GC u Shewhart 8 (ver Figura 31 y Anexo 8) se observó que todas tenían sesgo negativo cuya severidad variaba, para u constante, en función del tamaño de muestra (n). Ahora bien, el hecho de que todas las curvas analizadas tuviesen sesgo negativo, parecía indicar que con el GC u Shewhart no era posible construir gráficos de control capaces de detectar, de una forma idónea, mejoras en el proceso. Para determinar con certeza si esto era o no verdad, se decidió caracterizar el comportamiento del sesgo de las curvas ARL provenientes de gráficos u Shewhart.

Cabe mencionar que el estudio de caracterización se realizó luego de haber obtenido los resultados expuestos en las secciones 5.2 a 5.5, por lo que para este caso se utilizó directamente el parámetro ARL_{BSL} definido en la sección 5.4. Los resultados del estudio de caracterización se encuentran en la sección 5.6.1.

4.2.7 Sinopsis de la sección 4.2

La sección 4.2 resume exclusivamente el marco teórico que sirvió de base para determinar el efecto que tiene en la capacidad de monitorización de GCs tipo Shewhart, el inexacto ajuste de la aproximación normal a la distribución de Poisson. Todo el trabajo de investigación relevante y posterior a lo presentado en esta sección se incluye en la sección 5.6.

5 *Resultados y contribuciones al conocimiento*

5.1 **Caracterización y análisis de las colas alfa y el ARL_0**

En esta sección se destacan de una forma breve, un método de análisis y otros aspectos novedosos que fueron desarrollados durante la fase exploratoria de la investigación y que constan en la sección 4. Se los menciona en la sección 5, aún a costa de cierta reiteración, por tratarse de desarrollos originales que forman parte de los resultados obtenidos de la investigación.

5.1.1 *Método gráfico para analizar el comportamiento de las colas alfa*

Lo expuesto en las secciones 4.1.6 y 4.2.5 demuestra claramente que los valores de las colas alfa son de fundamental importancia para los gráficos de control (GCs) basados en las distribuciones binomial y de Poisson, ya que estas afectan directamente a su capacidad de monitorización.

Para caracterizar el comportamiento de las colas alfa, en la fase exploratoria de la investigación se desarrolló un algoritmo capaz de computar estos valores, y se utilizó un método gráfico para presentar los resultados obtenidos, para p fijo, acorde a un rango de tamaños de muestras. Ejemplos de estos se dan en las secciones 4.1.6.1 (Figura 17) y 4.2.5.1 (Figura 29). El concepto de “rango de tamaños de muestra” consta en la sección 4.1.6.1.

El analizar los resultados mediante un rango de tamaños de muestra, permitió desvelar varios aspectos del comportamiento de las colas alfa que eran desconocidos, como por ejemplo su oscilación y cuan alejadas se encontraban del valor nominal 0.00135. Cabe destacar que el método gráfico antes mencionado fue inicialmente desarrollado para analizar las colas alfa del GC p Shewhart, pero que al final fue empleado para analizar las colas alfa de todos los GCs incluidos en esta tesis, y que sirvió tanto para compararlos, así como para desarrollar nuevos gráficos.

Los resultados obtenidos en esta tesis ponen de manifiesto el potencial del análisis gráfico de las colas alfa. Este método podría fácilmente ser utilizado para analizar gráficos de control que están fuera del alcance de esta tesis.

5.1.2 Método gráfico para analizar el comportamiento del ARL_0

Con base en lo expuesto en la sección 4, es posible afirmar que uno de los parámetros que determinan la idoneidad de un GC para monitorizar procesos, es el ARL_0 . Ahora bien, en la fase exploratoria de la investigación se desarrolló un algoritmo capaz de computar el valor de este parámetro, y al igual que para las colas alfa, se utilizó un método gráfico para presentar los resultados obtenidos acorde a un rango de tamaños de muestras. Ejemplos de estos gráficos se hallan en las secciones 4.1.6.2 (Figura 18) y 4.2.5.2 (Figura 30).

El analizar los resultados de esta manera, hizo posible descubrir características del comportamiento del ARL_0 que eran desconocidas, y que hubiesen sido difíciles de revelar de otra forma, como la manera en que oscila y su amplitud. Cabe mencionar que es una práctica común simplemente asumir que el ARL_0 tiene un valor igual a 370, algo que como se ha comprobado, es totalmente erróneo.

Comprender que el ARL_0 no era constante y que más bien éste variaba y fluctuaba, además de como lo hacía, nos llevó a la conclusión de que el ARL_0 debería ser siempre computado y también a proponer un rango dentro del cual el ARL_0 podría ser considerado como aceptable. Esto facilitó el análisis y comparación de todos los GCs incluidos en esta tesis, y también sirvió para desarrollar nuevos gráficos, ver sección 5.1.3.

El potencial del análisis gráfico del ARL_0 es evidente, y podría ser fácilmente aplicado para analizar otros tipos de GCs que están fuera del alcance de esta tesis.

5.1.3 Determinación del rango de ARL_0 aceptables

En la sección 4.1.6.2 se demostró que el valor del ARL_0 para el GC p Shewhart no es igual a 370, como generalmente se presume, si no que por el contrario varía y fluctúa. Esto nos llevó a la conclusión de que no es correcto asumir que el ARL_0 de un GC es siempre 370, y que más bien su valor debería ser computado y evaluado como aceptable o no, según un rango de ARL_0 previamente establecido.

Ahora bien, para poder establecer un rango de ARL_0 aceptables, se debía tomar en cuenta que tanto valores muy altos como muy bajos de ARL_0 son indeseables para una idónea monitorización del proceso, esto debido a que un ARL_0 muy alto afecta la rapidez del gráfico para detectar cambios en el proceso (por su impacto sobre el ARL_1), mientras que un ARL_0 muy bajo da lugar a demasiadas falsas alarmas (Qiu, 2014, p. 80). El rango de ARL_0 aceptables que fue utilizado a lo largo del trabajo de investigación fue $250 \leq ARL_0 \leq 450$.

El rango $250 \leq ARL_0 \leq 450$ fue determinado de una manera empírica tomando en cuenta los siguientes criterios:

1. La variación del ARL_0 en función de la probabilidad de error tipo I (α).

La Figura 32 fue obtenida utilizando la ecuación (8) y muestra la variación del ARL_0 en función de α . Como se puede apreciar cambios pequeños de α por debajo de 0.0027, causan variaciones en el ARL_0 mucho más pronunciadas que cambios de igual magnitud por encima de 0.0027. Este comportamiento aplica para cualquier GC cuyos límites estén basados en la probabilidad de cometer un error de tipo I (ver sección 4.1.4.2).

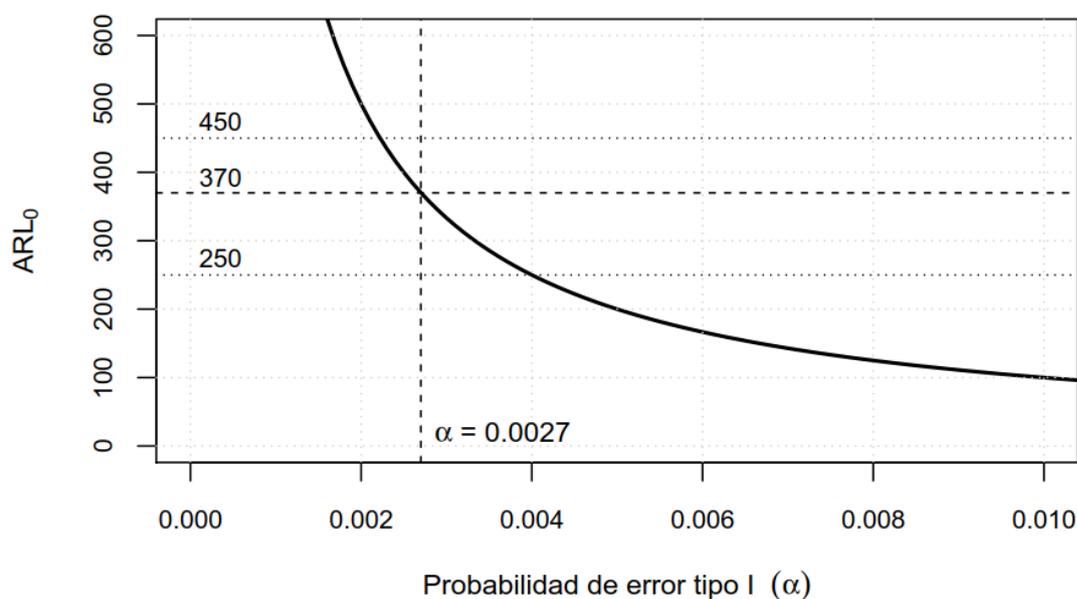


Figura 32. Relación entre la probabilidad de error Tipo I (α) y el ARL_0

2. La oscilación del ARL_0 expuesta en la sección 4.1.6.2.

La Figura 18 muestra un ejemplo de la variación típica del ARL_0 en GCs p Shewhart. En esa figura se puede apreciar claramente que la variación promedio se encuentra por debajo de 370, en ese ejemplo el promedio es 334. En general, el ARL_0 promedio que se obtiene con las colas alfa computadas en la sección 4.1.6.1 está por debajo de 370. Esto nos llevó a concluir que el rango de ARL_0 aceptables que se propusiese, tal vez no debería estar centrado en 370, si no un poco más abajo.

Con base en los dos aspectos antes descritos, se decidió, como propuesta de trabajo, que el rango de ARL_0 aceptables fuese 350 ± 100 , es decir $250 \leq ARL_0 \leq 450$.

Cabe mencionar que el rango $250 \leq ARL_0 \leq 450$ fue en un principio concebido para analizar el ARL_0 del GC p Shewhart, pero que a medida que avanzaba la investigación se razonó que

podría ser utilizado para cualquier otro GC por atributos. Cabe destacar que ese rango está incluido en la publicación más destacada que se ha obtenido para este trabajo de investigación, por lo que se asume que los expertos que revisaron el manuscrito estuvieron de acuerdo con su uso, ver Argoti& Carrión-García (2019a).

El rango de ARL_0 aceptables para GCs por atributos, es una propuesta novedosa que podría ser adaptada a otros GCs no incluidos en esta tesis.

5.1.4 El ratio de las colas alfa (R_α)

Un nuevo parámetro que se presenta en esta tesis es el “ratio de las colas alfa (R_α)”, este parámetro fue concebido en la fase exploratoria de la investigación y su descripción se encuentra en la sección 4.1.7.1.

El R_α fue utilizado para derivar una ecuación que sirve para estimar la severidad del sesgo de las curvas ARL para GCs p Shewhart, ver sección 5.3.1, así como para desarrollar un método muy simple, mediante el cual es posible verificar si los nuevos gráficos de control propuestos en esta tesis, tienen curvas ARL cuasi insesgadas, sin la necesidad de computar los parámetros de sesgo de la curva. Esta técnica se detalla en las secciones 5.5.3.4 y 5.6.3.4.

El R_α es un parámetro que podría ser utilizado en otros tipos de GCs que están fuera del alcance de esta tesis.

5.2 Gráfico de control p Shewhart: Caracterización del ARL-sesgo(%) y R_α en función de p y n

En esta sección se resumen los resultados del estudio realizado para determinar la relación que existe entre el ARL-sesgo(%), y el ratio de las colas alfa (R_α) en GCs p Shewhart. Los antecedentes del por qué se realizó este estudio se exponen en la sección 4.1.8.2. Cabe recordar también que las definiciones de los parámetros ARL-sesgo(%) y R_α se encuentran en la sección 4.1.8.1.

Para realizar el estudio se creó un algoritmo en R capaz de computar los valores del ARL-sesgo(%) y R_α para cualquier valor de p y n . Los cálculos fueron hechos para valores de fracción de no conformes (p) entre 0.01(0.01)0.20 y para rangos de muestra (n) específicos para cada p y que fueron determinados siguiendo los mismos lineamientos utilizados para la caracterización de las colas alfa en la sección 4.1.6.1. En esta tesis por razones de concisión se presentan solo los resultados para $p = 0.1$ y 0.05, los resultados para todos los otros p fueron muy similares por lo que se consideró innecesarios incluirlos.

La Figura 33 y la Figura 34 muestran resultados típicos de la variación del R_α y ARL-sesgo(%), otros ejemplos se incluyen en el Anexo 4. Como se puede observar la variación del R_α es muy pronunciada, con saltos muy significativos entre los puntos mínimos y máximos. Esto pone en evidencia la considerable fluctuación que pueden tener las colas alfa del GC p Shewhart con tan solo variar unas pocas unidades de n .

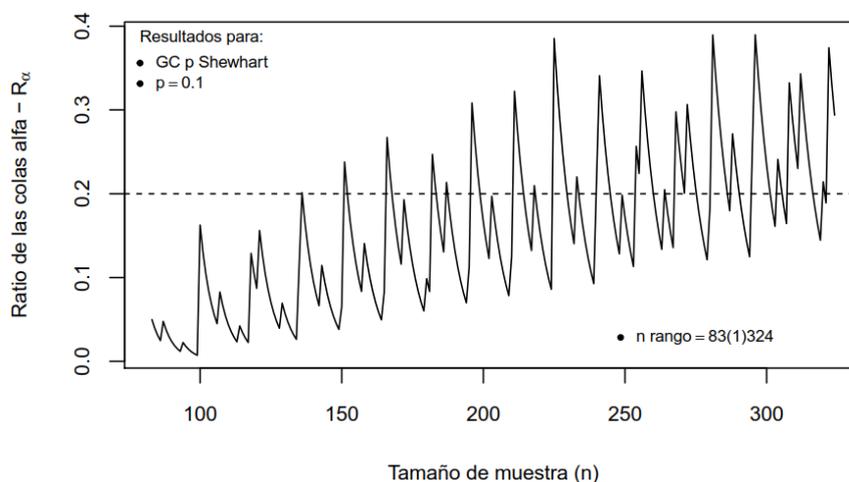


Figura 33. GC p Shewhart: Variación del R_α en función de n para $p = 0.1$

Algo importante que dejan de manifiesto los resultados, es que los valores de R_α , para un mismo p , pueden repetirse para diferentes tamaños de muestra. Por ejemplo en la Figura 33 se observa que el valor $R_\alpha = 0.2$ podría ser alcanzado por varios tamaños de muestra entre 136 y 316. Ahora bien, si observamos en la Figura 34 lo que sucede con el valor del ARL-sesgo(%) dentro de ese mismo rango de muestras, se puede concluir que este varía, por lo que GCs con un mismo valor de R_α podrían tener diferentes valores de ARL-sesgo(%). Esto conduce a la conclusión de que la severidad del sesgo en las curvas ARL, **no** es directamente proporcional al ratio de las colas alfa como en un principio se había intuido.

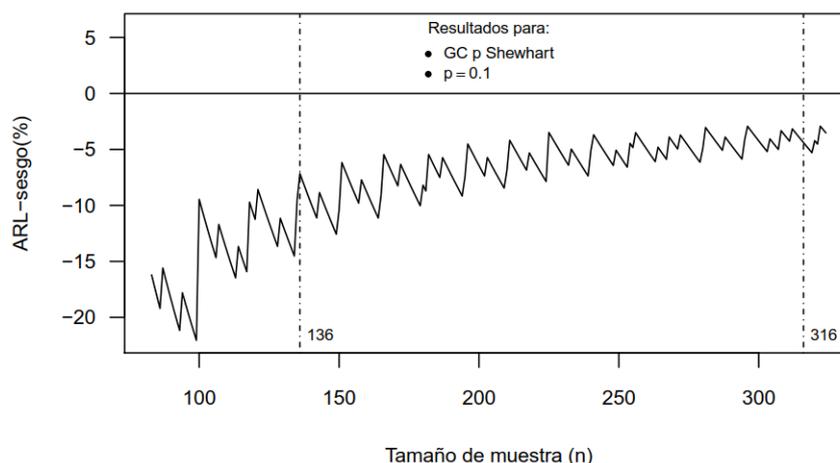


Figura 34. GC p Shewhart: Variación del ARL-sesgo(%) para $p = 0.1$

Por otra parte, analizando la Figura 34 y recordando que el ARL-sesgo(%) sirve para determinar la severidad del sesgo de las curvas ARL y que su valor ideal es cero (ver sección 4.1.8.1). Se concluye que, en términos generales, la severidad del sesgo oscila y varía de una forma logarítmica y que el nivel de severidad incrementa a medida que n disminuye y que además, como se había intuido inicialmente, el sesgo es siempre negativo, al menos para los valores de n utilizados en el estudio. Cabe recordar que un sesgo negativo implica que el GC tiene una reducida capacidad para detectar mejoras en el proceso, o en otras palabras, que tardara más en detectar reducciones de p (mejoras del proceso) que incrementos (deterioros del proceso).

Fue en este punto de la investigación cuando también se llegó a la conclusión de que para caracterizar el sesgo de las curvas ARL de una forma más apropiada, había que utilizar no solo el ARL-sesgo(%), sino también los parámetros ARL_0 y ARL_M ya que estos determinan la amplitud de la curva (ver Figura 23). Ese razonamiento condujo a la creación de los parámetros ARL-ratio y ARL_{BSL} que se presentan en la sección 5.4.

Por otro lado, la Figura 35 muestra un ejemplo de la relación característica que existe entre el ARL-sesgo(%) y el R_α . Como se puede apreciar el ARL-sesgo(%), en términos generales, tiende a reducir su valor de una forma logarítmica a medida que el R_α incrementa. En otras palabras, la severidad del sesgo de las curvas ARL para GCs p Shewhart, tiende a disminuir logarítmicamente a medida que la diferencia entre las colas alfa se reduce.

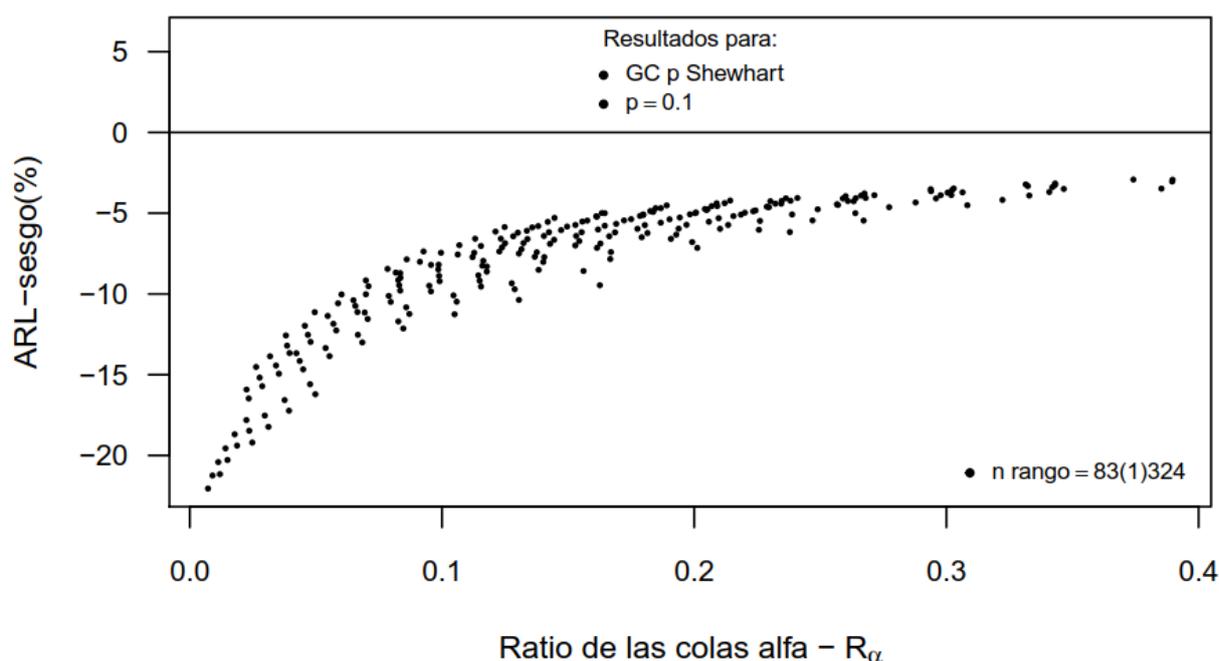


Figura 35. GC p Shewhart: Relación entre ARL-sesgo(%) y R_α para $p= 0.1$

Los resultados obtenidos en la fase de investigación que se resume en esta sección, sirvieron de base para el desarrollo de los métodos para estimar la severidad del sesgo que se detallan en la sección 5.3.

5.3 Gráficos de Control p Shewhart: Nuevos métodos para estimar la severidad del sesgo de las curvas ARL

En esta sección se resume el trabajo realizado con la finalidad de desarrollar métodos para estimar la severidad del sesgo en las curvas ARL del GC p Shewhart. Estos métodos fueron desarrollados debido a que el GC p Shewhart es muy popular, siendo de esencial enseñanza en cualquier curso de control estadístico de procesos, sin embargo y como se ha demostrado anteriormente (ver sección 5.2), este GC tiene la falencia de que su curva ARL tiene un sesgo negativo que en ocasiones puede ser muy severo.

Ahora bien, si alguien, con el fin de establecer la idoneidad de un GC p Shewhart, quisiera determinar cuan sesgada es la curva ARL, tendría que necesariamente obtener los valores de los parámetros ARL-sesgo (%), ARL_0 y ARL_M (ver Tabla 6). Y para hacer esto necesitaría computar los valores con los que se traza la curva ARL, lo que requiere de cálculos que pueden ser largos y complejos.

Lo expuesto en el párrafo anterior, llevó a la conclusión de que sería novedoso crear métodos que ayuden a estimar la severidad del sesgo de una manera simple y efectiva, sin que sea necesario computar y analizar la curva ARL. Cabe mencionar que estos métodos fueron desarrollados antes de que el parámetro ARL_{BSL} (ver sección 5.4) fuese concebido.

Los resultados parciales obtenidos de la investigación descrita en esta sección fueron presentados en un congreso internacional, ver Argoti&Carrión-García (2017) y los resultados totales dieron lugar a la primera publicación obtenida para esta tesis doctoral, ver Argoti&Carrión-García (2018a). Además con lo desarrollado en esta sección se cumplió con el objetivo específico número dos del trabajo de investigación.

5.3.1 Ecuación para estimar la severidad del sesgo en la curva ARL

El primer método desarrollado para estimar la severidad del sesgo hace uso de la ecuación (24) y de la Tabla 8. Para obtener esta ecuación primeramente se computaron los valores del ARL-sesgo(%) y R_α para fracción de no-conformes entre $p = 0.005(0.005)0.20$ usando rangos de tamaños de muestra específico para cada p , donde el n mínimo se lo determinó

usando $d = p/2$ en la ecuación (13) y el n máximo fue 10000 para todos los casos, un ejemplo de los resultados obtenidos (para $p = 0.05$) se da en la Figura 36.

A los resultados obtenidos en cada p se aplicó el modelo de regresión ARL- sesgo($\%$) = $a_p + b_p * \log(R_\alpha)$ obteniendo así una ecuación para estimar el valor del ARL- sesgo($\%$) en función de R_α para cada uno de valores de p utilizados. Cabe mencionar que los valores de los coeficientes a_p y b_p fueron específicos para cada p utilizado.

Ahora bien, algo que se identificó fue que cuando los valores de los coeficientes a_p y b_p se los trazaba en función de p , estos seguían una distribución logarítmica bien definida. Este hecho condujo a aplicar el modelo de regresión lineal-log a estos valores dando como resultado las siguientes ecuaciones i) $a_p = -0.89 - 0.87 * \log(p)$ y ii) $b_p = 2.63 - 0.52 * \log(p)$. Nótese que reemplazando las ecuaciones i) y ii) en la formula ARL- sesgo($\%$) = $a_p + b_p * \log(R_\alpha)$ se obtiene la ecuación (24), con la que es posible estimar el valor del ARL- sesgo($\%$) para cualquier valor de p (entre 0.005 y 0.2) y n .

Por otra parte, además de lo mencionado anteriormente, los criterios de la Tabla 8 fueron establecidos luego de un exhaustivo análisis de esos resultados del ARL- sesgo($\%$) y R_α . Cabe recordar que la definición del ARL- sesgo($\%$) se da en la Tabla 6 y que el R_α es el ratio de las colas alfa.

$$ARL- sesgo(\%) \approx -0.89 - 0.87 \log_e(p) + \log_e(R_\alpha) [2.63 - 0.52 \log_e(p)] \quad (24)$$

Donde:

- $p =$ Fracción de no- conformes del proceso
- $R_\alpha =$ Ratio de las colas alfa

La Figura 36 muestra un ejemplo del ajuste de los valores ARL- sesgo($\%$) predichos por la ecuación (24) para $p = 0.05$, como se puede apreciar el ajuste es aceptable. Cabe destacar que la ecuación da resultados razonables desde $p = 0.006$ hasta $p = 0.20$. Sin embargo, en algunos casos los valores predichos por (24) podrían estar por encima de cero, si esto sucediera la recomendación que se da es que el nivel de severidad para estos caso sea: i) Leve: si $0 < ARL- sesgo(\%) \leq 1$ y ii) Insignificante: si $ARL- sesgo(\%) > 1$.

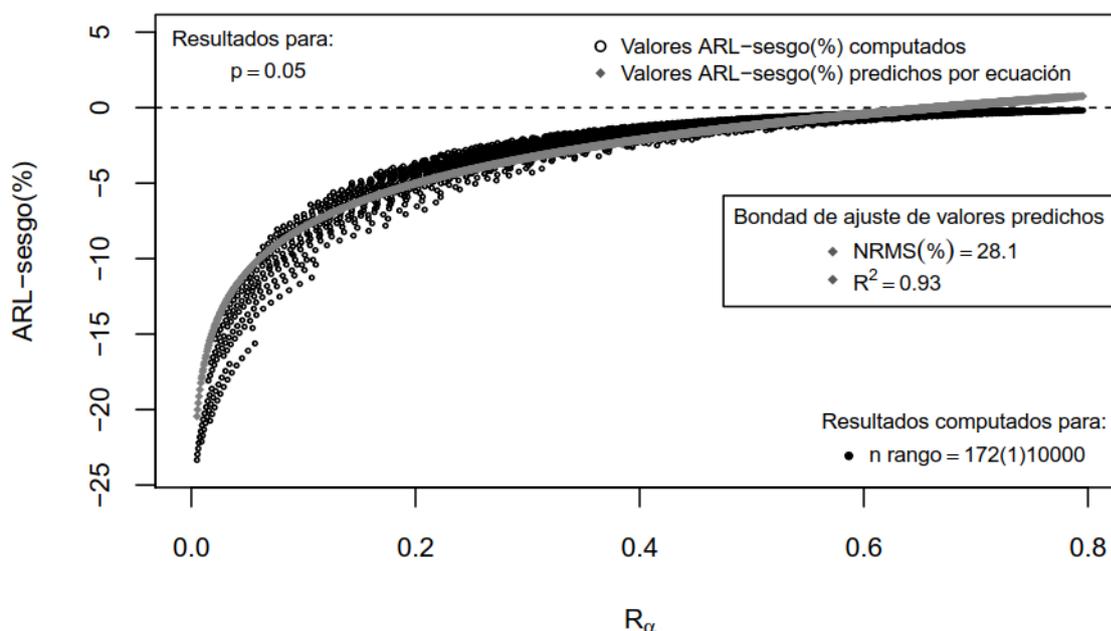


Figura 36: Bondad de ajuste de los valores ARL-sesgo(%) predichos por la ecuación (24)

Como se puede apreciar, con la ecuación (24) solo se requieren p y R_α para obtener una estimación del ARL-sesgo(%). Recordemos que para obtener los valores de las colas alfa es necesario aplicar la probabilidad binomial acumulada, ver ecuaciones (10) y (11). Ahora bien, la probabilidad binomial acumulada viene incorporada en varios programas computacionales de uso común, como Excel, por lo que obtener esos valores es bastante sencillo.

Una vez obtenido el valor del ARL-sesgo(%), se debe utilizar la Tabla 8 para establecer el nivel de severidad del sesgo. Por ejemplo para un GC con $p = 0.02$ y $n = 600$ se obtiene un $R_\alpha = 0.03$, aplicando estos valores en (24) se obtiene que el ARL-sesgo(%) ≈ -13.8 , por lo que el nivel de severidad de la curva ARL será considerable. La recomendación que se dio en la publicación correspondiente y que se ratifica en esta tesis, es que solo GCs con curvas ARL que se hallen en los criterios leve e insignificante sean consideradas como adecuadas para monitorizar procesos.

Por otra parte, mediante las ecuaciones (25) y (12) también es posible obtener un valor estimado del ARL_M , para el ejemplo del párrafo anterior mediante estas ecuaciones se obtiene que el ARL_M es 1339. Aparte de eso, con el valor de las colas alfa, se podría obtener el ARL_0 , que para el caso del ejemplo antes mencionado es 354.

$$p_M = [p + p(\text{ARL-sesgo}(\%) / 100)] \quad (25)$$

Dónde:

$P_M = p$ -modificado donde sucede el ARL_M

Tabla 8: Criterios para determinar la severidad del sesgo en función del ARL-sesgo(%)

<i>Nivel de severidad</i>	<i>ARL-sesgo(%)</i>
1. Considerable	$ARL\text{-sesgo}(\%) < -10\%$
2. Significativo	$-10\% \leq ARL\text{-sesgo}(\%) < -3\%$
3. Moderado	$-3\% \leq ARL\text{-sesgo}(\%) < -1\%$
4. Leve	$-1\% \leq ARL\text{-sesgo}(\%) < -0.5\%$
5. Insignificante	$-0.5\% \leq ARL\text{-sesgo}(\%) \leq 0\%$

Un aspecto que se tomó en cuenta para desarrollar la Tabla 8 es la relación que existe entre los parámetros ARL_0 , ARL_M y $ARL\text{-sesgo}(\%)$, recordemos que estos tres parámetros definen la severidad del sesgo en las curvas ARL (ver sección 4.1.8.1). El análisis para establecer esta relación se lo realizó para $0.006 \leq p \leq 0.2$. En este punto de la investigación, y con el objetivo de reducir, de tres a dos, los parámetros con los que se define el sesgo, se concibió el ARL-ratio que no es más que el cociente de ARL_M / ARL_0 . Todos los gráficos de curvas ARL en esta tesis incluyen el valor de este parámetro, así por ejemplo en la Figura 23 (ver sección 4.1.8) el ARL-ratio es igual a 1.36, lo que indica que el ARL_M es 1.36 veces más grande que el ARL_0 . Cabe mencionar que lo óptimo sería que el ARL_M y el ARL_0 tengan el mismo valor, o en otras palabras que el ARL-ratio sea igual a uno ($ARL\text{-ratio} = 1$).

Un ejemplo de los resultados y del análisis hecho para determinar la relación entre el ARL-ratio y el $ARL\text{-sesgo}(\%)$ se presenta en la Figura 37, como se puede apreciar la relación que existe es logarítmica. En este punto es preciso mencionar que la curva ARL “ideal” es **insesgada** y que esta se da solo cuando se tienen valores de $ARL\text{-ratio} = 1$ y $ARL\text{-sesgo}(\%) = 0$. Cabe recordar que una curva ARL insesgada denota que el GC al cual pertenece, puede detectar de igual manera tanto mejoras como deterioros del proceso.

Ahora bien, lo que muestra la Figura 37 revela que los GCs p Shewhart pueden llegar a tener curvas ARL con valores de ARL-ratio y $ARL\text{-sesgo}(\%)$ extremadamente alejados de lo óptimo. El comportamiento del $ARL\text{-sesgo}(\%)$ vs ARL-ratio fue similar para todos los otros valores de p utilizados en el análisis.

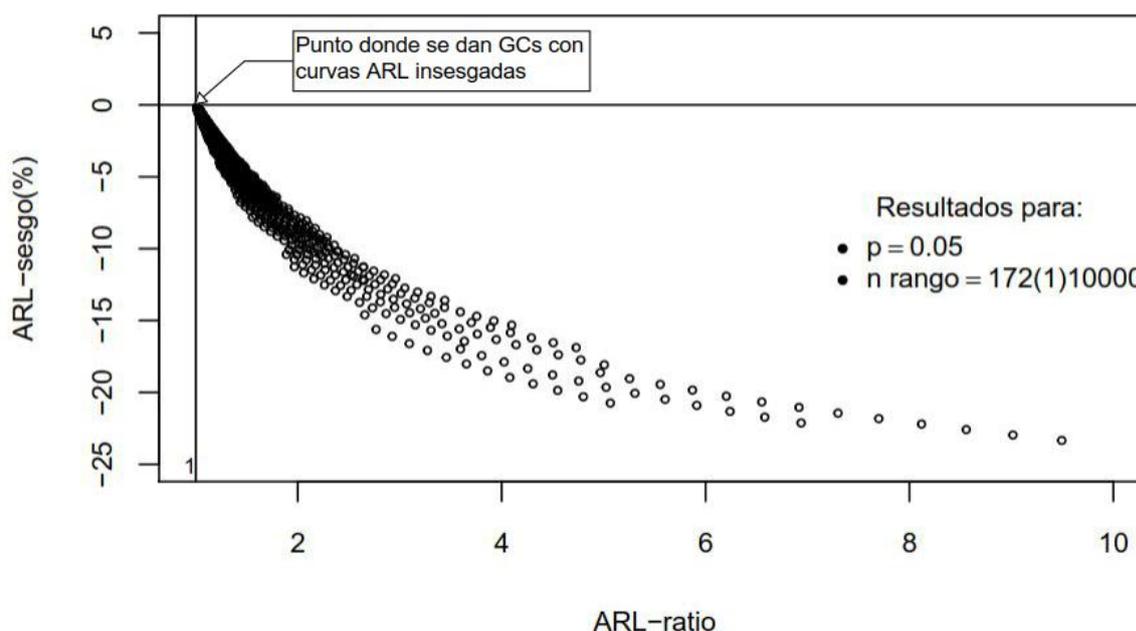


Figura 37: Ejemplo de la relación entre ARL-sesgo(%) y ARL-ratio para $p = 0.05$

5.3.1.1 Curvas ARL cuasi incesgadas: definición

Algo importante a destacar es que durante la fase de investigación resumida en la sección 5.3.1, se concibió por primera vez el concepto de **curvas ARL cuasi incesgadas**, a las que se definió como: “*curvas ARL cuyo sesgo tiene un impacto insignificante en la capacidad de monitorización de gráficos de control*”. Además, en un principio se definieron a estas como curvas cuyos parámetros de sesgo debían estar dentro de los valores: i) ARL-sesgo(%) $> -1\%$ y ii) ARL-ratio < 1.1 . Por eso la recomendación que se da en la sección 5.3.1, y en la publicación correspondiente, de que solo GCs p Shewhart que tengan curvas ARL cuya severidad de sesgo este entre leve e insignificante, sean considerados como óptimos para monitorizar procesos.

5.3.2 Zonas de severidad del sesgo para curvas ARL

EL segundo método desarrollado para facilitar la estimación de la severidad del sesgo en las curvas ARL del GC p Shewhart, hace uso de zonas severidad. Estas zonas se muestran en la Figura 39.

Para crear estas zonas, se utilizaron los resultados del ARL-sesgo(%) obtenidos en la sección 5.3.1. Estos valores fueron clasificados en tres categorías de acuerdo a n y a su valor ARL-sesgo(%) correspondiente. Un ejemplo de esta clasificación, para $p = 0.05$, se muestra en la Figura 38.

El criterio para obtener los límites de cada categoría se exponen en la Tabla 9 donde, y a modo de ejemplo, también se incluyen los límites para el caso de $p = 0.05$. Mediante estos criterios se logró establecer los límites de las zonas de severidad que se muestran en la Figura 39. La Tabla 10 muestra una selección de estos límites con los que se podría fácilmente trazar esas zonas.

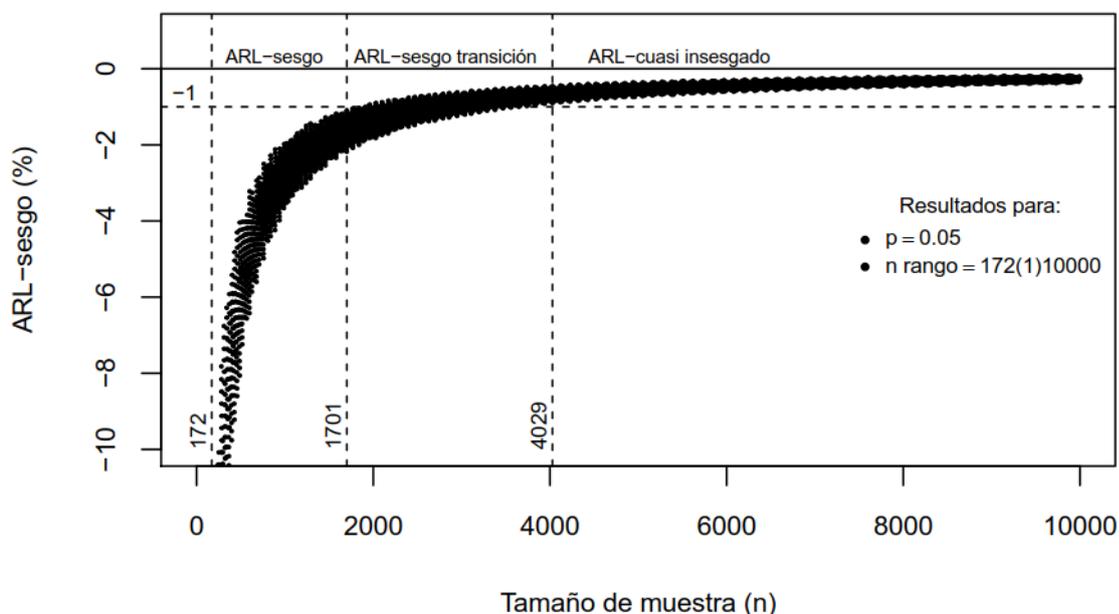


Figura 38. Gráfico de control p Shewhart: ARL-sesgo(%) para $p = 0.05$

Tabla 9. Criterios para establecer los límites de las zonas de severidad

Categoría	Criterios que determinan el rango de n de cada categoría	Límites n de cada categoría para $p = 0.05$	
		Limite n inferior	Limite n superior
ARL-sesgo	<p>En esta categoría se agrupan los GCs cuyas curvas ARL tienen un ARL-sesgo(%) $< -1\%$. Por lo que la severidad del sesgo de esas curvas podría estar dentro de las categorías 1, 2 y 3 de la Tabla 8.</p> <p>En general, mientras más cerca n esté del límite superior, menor será la severidad del sesgo.</p> <p>El límite inferior está dado por el tamaño de muestra mínimo debajo del cual ya no se obtienen límites de control inferior (LCI), y que se obtiene mediante la ecuación (13). Cabe destacar que una curva ARL sesgada se da solo en GCs que tienen límites de control inferior (LCI) y superior (LCS).</p>	172	1701

	Razón por la cual esta tesis solo trata GCs que tienen esos dos límites.		
ARL-sesgo transición	En esta categoría se agrupan los GCs cuyas curvas ARL podrían tener $-3 \leq \text{ARL-sesgo}(\%) < 0$. Por lo que la severidad del sesgo de esas curvas podrían caer dentro de las categorías 3, 4 y 5 de la Tabla 8. En general, mientras más cerca n esté del límite superior, más se acerca estará la curva de la categoría 5.	1702	4029
ARL-cuasi insesgado	En esta categoría se agrupan los GCs cuyas curvas ARL tienen $-1 < \text{ARL-sesgo}(\%) \leq 0$. Por lo que esas curvas podrían caer dentro de las categorías 4 y 5 de la Tabla 8.	4030	10000 o más

Tabla 10. Ejemplo de límites n utilizado para trazar las zonas de severidad

	p	0.01	0.02	0.04	0.05	0.06	0.08	0.1	0.12	0.14
ARL sesgo	n inferior	892	442	217	172	142	105	82	67	56
	n superior	7003	3976	2188	1701	1434	1026	806	609	468
ARL sesgo transición	n inferior	7004	3977	2189	1702	1435	1027	807	610	469
	n superior	38447	12547	5236	4029	3224	2243	1689	1103	956
ARL cuasi insesgado	n inferior	38448	12548	5237	4030	3225	2244	1690	1104	957
	n superior		10000 o más							

La Figura 39 permite estimar la severidad del sesgo en las curvas ARL de los gráficos p Shewhart, de una manera rápida y sencilla, ya que solo se requiere identificar la zona en que la que se ubican los parámetros p y n . Por ejemplo si un GC tienen parámetros $p = 0.04$ y $n = 1000$ entonces su curva ARL tendría un sesgo demasiado severo. En cambio con $p = 0.04$ y $n = 2500$ el sesgo podría estar dentro de los niveles de severidad moderada a insignificante. Ahora bien, la recomendación que se da en caso de que se requiera establecer el nivel de severidad específico es que se utilice la ecuación (24).

Cabe mencionar que información complementaria que podría ser de interés acerca de los métodos para estimar la severidad del sesgo en las curvas ARL, se da en la publicación correspondiente, ver Argoti&Carrión-García (2018a).

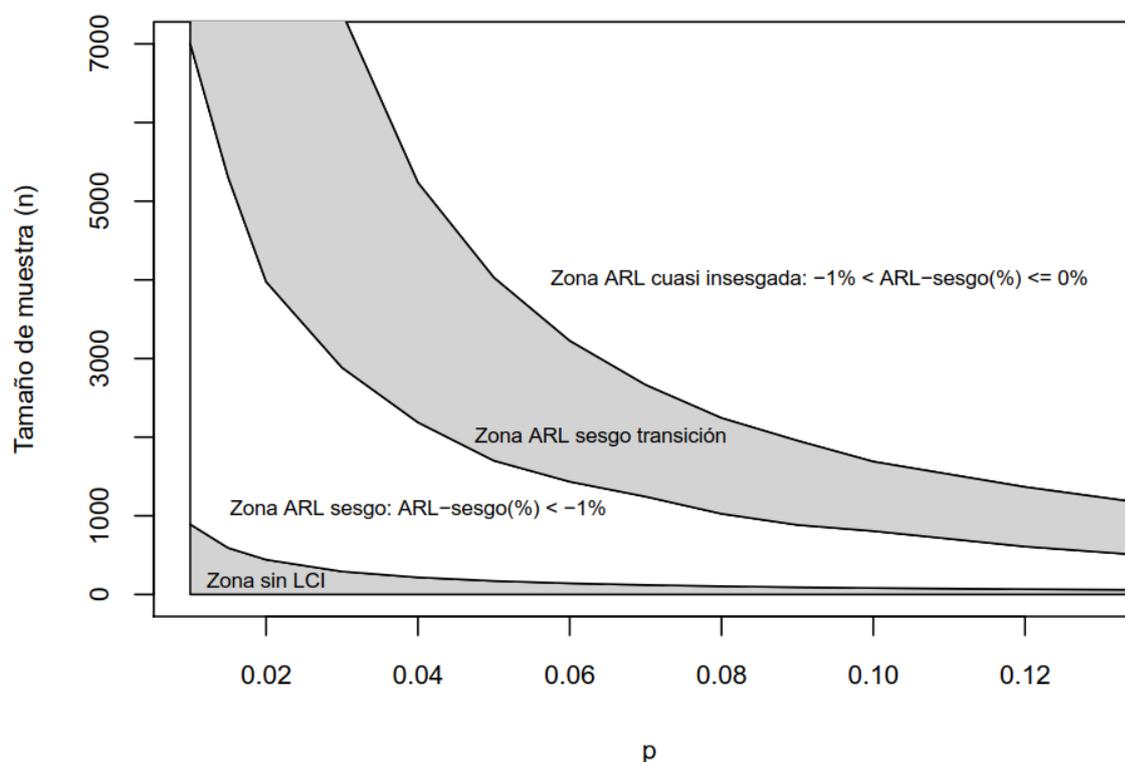


Figura 39: Zonas de severidad del sesgo para curvas ARL de GCs p Shewhart

5.4 El ARL_{BSL} (ARL bias severity level) o nivel de severidad del sesgo en la curva ARL

5.4.1 El parámetro ARL_{BSL} : Definición y criterios para interpretarlo

En la sección 5.2 se resume el primer trabajo realizado con la finalidad de caracterizar la severidad del sesgo en las curvas ARL, allí se menciona que fue en esa fase de la investigación que se llegó a la conclusión de que, para caracterizar el sesgo de las curvas ARL de una forma más apropiada, había que utilizar no solo el $ARL\text{-sesgo}(\%)$, sino también los parámetros ARL_0 y ARL_M ya que estos determinan la amplitud de la curva, ver Figura 23. Por esta razón, y luego de un extenso proceso de análisis, se concibió el ARL_{BSL} que fusiona esos tres parámetros y permite determinar la severidad del sesgo de una forma efectiva. El ARL_{BSL} se computa mediante la ecuación (26). Los criterios para interpretar los valores que se obtienen con esa ecuación se dan en la Tabla 11.

$$ARL_{BSL} = ARL\text{-sesgo}(\%) * ARL\text{-ratio} \quad (26)$$

Dónde:

- El $ARL\text{-ratio} = ARL_M / ARL_0$

- Idealmente el $(ARL_M/ARL_0) = 1$, caso contrario mientras más grande sea el ratio mayor será la severidad del sesgo.
- $ARL\text{-sesgo}(\%) = 100 * [(p_M/p) - 1]$

$p_M = p$ -modificado donde sucede el ARL_M .
 $p =$ fracción de no conformes del proceso bajo control

- El $ARL\text{-sesgo}(\%)$, también llamado sesgo ARL relativo, es la diferencia proporcional entre la ubicación del ARL_M y el ARL_0 en el eje- x , ver Figura 23.
- Idealmente el $ARL\text{-sesgo}(\%) = 0$ debería ser igual a cero. Mientras más se aleje de cero mayor será la severidad del sesgo.
- Un $ARL\text{-sesgo}(\%)$ negativo indica que la curva tiene sesgo negativo.

Tabla 11: Criterios para la interpretación del ARL_{BSL}

- i. $ARL_{BSL} = 0$, indica que la curva es insesgada (esta es la condición ideal)
- ii. $ARL_{BSL} < 0$, indica que la curva tiene sesgo negativo.
- iii. $ARL_{BSL} > 0$, indica que la curva tiene sesgo positivo.
- iv. Mientras más se aleje el ARL_{BSL} de cero mayor será la severidad del sesgo.

El ARL_{BSL} hizo posible caracterizar la severidad del sesgo en las curvas ARL de todos los GCs incluidos en esta tesis. Lo que a su vez permitió realizar comparaciones entre gráficos existentes así como el desarrollo de nuevos gráficos. Cabe destacar que con el ARL_{BSL} se logró cumplir el objetivo específico número uno de la investigación.

Vale la pena señalar que el parámetro ARL_{BSL} podría fácilmente ser utilizado para analizar gráficos de control que están fuera del alcance de esta tesis.

5.4.2 Criterio ARL_{BSL} para obtener curvas ARL cuasi insesgadas

En la Tabla 11 se destaca que el ARL_{BSL} ideal es cero, ya que esto indica que la curva ARL es insesgada. Sin embargo, las curvas ARL de todos los GCs incluidos en este trabajo de investigación casi nunca llegan a obtener ese valor, por lo que éstas casi siempre serán sesgadas. Sin embargo, frecuentemente la severidad del sesgo es tan pequeña, que su impacto en la capacidad de monitorización del GC es insignificante. A este tipo de curvas se las definió como “cuasi insesgadas” en la sección 5.3.1.1.

Para determinar un rango de valores ARL_{BSL} que permitiesen identificar si una curva ARL era o no “cuasi insesgada”, se recurrió a los resultados del trabajo que se resume en la sección 5.3.1, en donde se determinó que curvas “cuasi insesgadas” se dan solo cuando los parámetros de sesgo están dentro de los valores: i) $ARL\text{-sesgo}(\%) > -1\%$ y ii) $ARL\text{-ratio} < 1.1$. Sin embargo, una falencia del criterio -1% para el $ARL\text{-sesgo}(\%)$ es que solo toma en cuenta curvas con sesgo negativo, cuando en la realidad las curvas ARL pueden tener tanto sesgo negativo como positivo, por lo que se llegó a la conclusión de que el criterio para el $ARL\text{-sesgo}(\%)$ debería ser $\pm 1\%$.

Ahora bien, utilizando la ecuación (25) con $ARL\text{-sesgo}(\%) > \pm 1\%$ y $ARL\text{-ratio} < 1.1$ se obtuvo el primer criterio ARL_{BSL} para obtener curvas cuasi insesgadas, que fue $-1.1 < ARL_{BSL} < 1.1$. Sin embargo, luego de llevar a cabo los primeros análisis con este criterio, se concluyó que era excesivamente exigente, por lo que luego de una cuidadosa valoración se determinó que era mejor extenderlo a $-2 < ARL_{BSL} < 2$.

El criterio $-2 < ARL_{BSL} < 2$ para curvas cuasi insesgadas fue aplicado a todos los GCs incluidos en esta tesis.

5.5 El gráfico de control p Kmod: Un nuevo gráfico para monitorizar procesos con base en la distribución binomial

En la sección 5.5 se agrupan resultados relevantes que llevaron a la consecución de la segunda publicación obtenida para esta tesis doctoral, ver Argoti&Carrión-García (2019a).

5.5.1 Gráfico de control p Shewhart: Caracterización del ARL_{BSL}

Con la finalidad de caracterizar el comportamiento del sesgo en las curvas ARL del GC p Shewhart, se desarrolló un algoritmo capaz de computar el ARL_{BSL} para cualquier valor de p y n . Con este algoritmo se realizó un extenso análisis del ARL_{BSL} para $p = 0.01(0.01)0.20$ con rangos de tamaños de muestra (n) específicos para cada p . Además también se calcularon los ratios de las colas alfa R_α .

Para determinar el rango de n , se aplicaron los siguientes criterios: i) que con todas las combinaciones de n y p se debían obtener límites de control inferior (LCI) y superior (LCS) y ii) que el tamaño de muestra debía ser lo suficientemente grande. Para cumplir estos criterios se utilizó la ecuación (13) con $d = 0.99p$ para el valor n mínimo y $d = p/2$ para el n máximo.

Ejemplos de los resultados se muestran en la Figura 40 y también en las figuras incluidas en el Anexo 5. En todas las figuras el rango de n se denota como n rango = n mínimo (1) n máximo, por ejemplo en la Figura 40 se muestra el n rango = 174(1)684, lo que indica que los cálculos fueron realizados desde el tamaño de muestra 174 hasta el 684 en incrementos de una unidad.

Tomando en cuenta los criterios de la Tabla 11, de los resultados obtenidos se concluye lo siguiente:

1. Que el GC p Shewhart produce gráficos de control con ARL_{BSL} negativos, lo que denota que las curvas ARL de estos gráficos tendrán sesgo negativo, lo que a su vez implica que estos GCs tendrán una reducida capacidad para detectar mejoras en el proceso.
2. Que el ARL_{BSL} en los GCs p Shewhart fluctúa por debajo de cero y que su valor específico depende de los parámetros p y n .
3. Que las curvas ARL de los GCs p Shewhart se acercan al umbral $ARL_{BSL} > -2$ (de curvas ARL cuasi-insesgadas) solo para tamaños de muestra extremadamente altos.

De lo expuesto en las conclusiones 1, 2 y 3 se deduce que con el gráfico de control p Shewhart, no es posible monitorizar procesos de una forma adecuada.

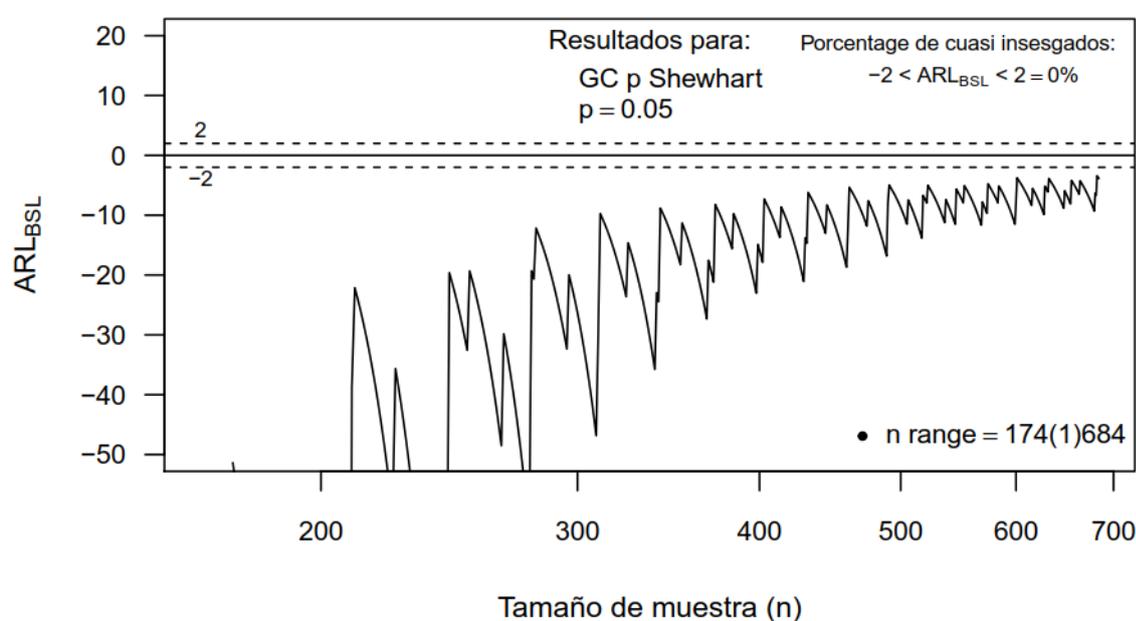


Figura 40. GC p Shewhart: Comportamiento del ARL_{BSL} para $p = 0.05$

5.5.1.1 GC p Shewhart: relación entre las colas alfa y el ARL_{BSL}

Utilizando los resultados obtenidos en el estudio descrito en la sección previa para $p = 0.05$, se creó la Figura 41 que muestra la relación que existe entre las colas alfa y el ARL_{BSL} . Cabe recalcar que aunque los resultados mostrados son para $p = 0.05$, el comportamiento es similar desde $p = 0.01$ hasta $p = 0.20$ (que es todo el rango de valores evaluado).

Antes de analizar los resultados en la Figura 41 hay que recordar que el ratio de las colas alfa es igual a $R_\alpha = \alpha_L/\alpha_U$ (ver sección 4.1.7.1) y que este sirve para determinar la disparidad que existe entre las colas alfa. Recordemos también que la situación ideal es que estas sean similares, es decir que $R_\alpha \approx 1$, y que mientras más se aleje el R_α de 1 mayor será la disparidad entre las colas.

Analizando los resultados expuestos es posible concluir lo siguiente:

1. En términos generales, cuanto más grande sea la disparidad de las colas alfa, mayor será la severidad del sesgo en las curvas ARL.
2. Que para obtener valores ARL_{BSL} cercanos a cero es necesario reducir, a lo mínimo posible, la disparidad entre las colas alfa.

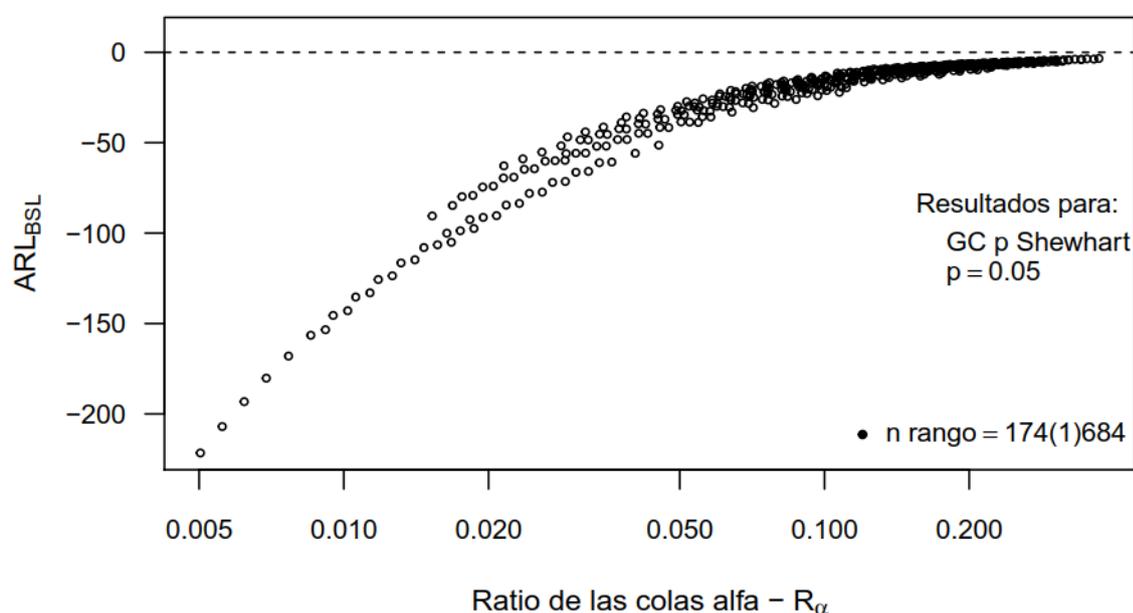


Figura 41. GC p Shewhart: Relación entre R_α y ARL_{BSL} para $p = 0.05$

5.5.1.2 Sinopsis de la sección 5.5.1

Al final de la sección 3.2.2.2 se mencionó que la pregunta que había dado lugar a la investigación resumida en esta tesis era: *¿Cómo afecta la imprecisión de la aproximación normal a la capacidad de monitorización de los CGs por atributos tipo Shewhart?*

Para obtener una respuesta a esa pregunta, para el caso de GCs por atributos basados en la distribución binomial, se desarrolló todo lo que consta en las secciones 4.1, 5.2, 5.3, 5.4 y 5.5.1. Por lo que finalmente, y fundamentado en lo expuesto en esas secciones, es posible afirmar que la imprecisión del ajuste de la aproximación normal a la distribución binomial, para $p < 0.2$, afecta de forma muy adversa la capacidad de monitorización del GC p Shewhart. Por este motivo la recomendación que se da en esta tesis es que este GC no sea utilizado, especialmente si se requiere detectar mejoras en el proceso.

5.5.2 Gráficos de Control p alternativos: Caracterización del ARL_0 y ARL_{BSL}

Luego de establecer que el GC p Shewhart tenía una capacidad de monitorización no óptima (ver sección 5.5.1), el trabajo de investigación cambió de enfoque y se encauzó en cumplir el tercer objetivo específico de la investigación, que fue: *Caracterizar la capacidad de monitorización de gráficos de control por atributos relevantes que hayan sido propuestos hasta la fecha de realización de la investigación.*

Para lograr ese objetivo, y para el caso específico de GCs basados en la distribución binomial, se efectuó una búsqueda de gráficos de control p que hubiesen sido propuestos como alternativas superiores al GC p Shewhart y que fuesen igual de fáciles de aplicar y utilizar. La búsqueda trajo a la luz varios GCs que sorprendentemente casi nunca son mencionados en textos de control estadístico.

Para crear estos “GCs p alternativos” sus autores utilizaron una gran gama de técnicas, razón por la cual los resultados de la búsqueda fueron clasificados en dos grupos de acuerdo al enfoque utilizado para crearlos. En esta sección se presentan estos grupos y también los resultados del análisis del ARL_0 y ARL_{BSL} para varios de estos GCs.

En este punto cabe mencionar que en esta tesis se considera que “*un gráfico de control es óptimo para monitorizar procesos, solo cuando su curva ARL es cuasi insesgada y su ARL_0 es aceptable*”. Se entiende que afirmar que un gráfico es óptimo, es igual a decir que su capacidad de monitorización es idónea.

Los análisis de todos los GCs alternativos fueron hechos siguiendo los mismos lineamientos, en cuanto al rango de p y n , que los utilizados en la sección 5.5.1 para el GC p Shewhart. Por otra parte, vale la pena recordar que el criterio ARL_{BSL} para establecer si una curva ARL es o no cuasi insesgada se da en la sección 5.4.2, mientras que el criterio para establecer si un ARL_0 es aceptable se encuentra en la sección 5.1.3.

5.5.2.1 Gráficos de control p : Enfoque optimización de los límites de control

En este grupo se han colocado los GCs que mediante ecuaciones simples, obtienen límites de control que dan valores de colas alfa cercanos a 0.00135. Los GCs que caen dentro de este enfoque se detallan en la Tabla 12.

Tabla 12: Gráficos de control p alternativos: Enfoque optimización de límites de control

<i>Descripción del GC</i>	<i>Autores y publicación</i>
GC p con límites computados mediante una ecuación obtenida a través de una regla empírica	Ryan T.P (1989, pág. 348)
GC p con límites computados mediante una ecuación basada en la expansión Corner-Fisher	Winterbottom (1993) Chen (1998)
GC p con límites computados a través de una ecuación basada en análisis de regresión.	Ryan y Schwertman (1997)
GC p con límites computados mediante una ecuación basada en el intervalo de Wilson.	Park (Park, 2013)

Los cuatro GCs incluidos en la Tabla 12 fueron sometidos a un estudio que tuvo como fin caracterizar el comportamiento de sus ARL_0 y ARL_{BSL} , y también determinar sus rendimientos en cuanto a la proporción de gráficos de control óptimos que se obtienen dentro de un rango de n . Los resultados de este estudio permitieron cumplir con el tercer objetivo específico de esta tesis.

Por motivos de concisión, se presentan solo los detalles y resultados de los gráficos más destacados, siendo estos: i) el GC p basado en la expansión Corner-Fisher y ii) el GC p basado en análisis de regresión.

5.5.2.1.1 Gráfico de control p basado en la expansión Corner-Fisher

Winterbottom (1993) fue quien originalmente propuso el GC p basado en la expansión Corner-Fisher. Años más tarde Chen (1998) lo volvió a presentar añadiendo variaciones a su aplicación. Los límites de control para este gráfico se calculan mediante la ecuación (27).

$$\text{Límites del GC } p \text{ Corner-Fisher} = p \pm K \sqrt{\frac{p(1-p)}{n} + \frac{4(1-2p)}{3n}} \quad (27)$$

Donde:

- Para el LC Inferior usar: +
- Para el LC Superior usar: -
- Para límites 3 sigma usar: $K=3$
- n = tamaño de muestra
- p = fracción promedio de no conformes del proceso
- Línea central del gráfico= p

La Figura 42 muestra un ejemplo de las colas alfa que se obtienen con los límites computados con (27). Comparando estas colas alfa con las del GC p Shewhart (ver Anexo 1), es fácil darse cuenta que en este caso las colas alfa oscilan alrededor de 0.00135 y no por encima o por debajo como es el caso para el GC Shewhart. En general el estudio de colas alfa del GC p Corner-Fisher demostró que estas, en promedio, si están más cerca de 0.00135 que las del GC p Shewhart.

Sin embargo, el hecho de que las colas alfa estén más cercanas a 0.00135, no es garantía de que los gráficos construidos con este GC tengan una capacidad de monitorización idónea, esto debido a que cada gráfico tendrá una combinación específica de parámetros binomiales p y n y por ende valores únicos de colas alfa, ARL_0 y ARL_{BSL} . Para establecer su verdadera capacidad de monitorización es necesario caracterizar su ARL_0 y ARL_{BSL} para una gran variedad de combinaciones p y n . Esto es justamente lo que se realizó, ejemplos de resultados para $p = 0.05$ se muestran en la Figura 43 y la Figura 44, además, en la Tabla 13 se sintetizan todos los resultados obtenidos.

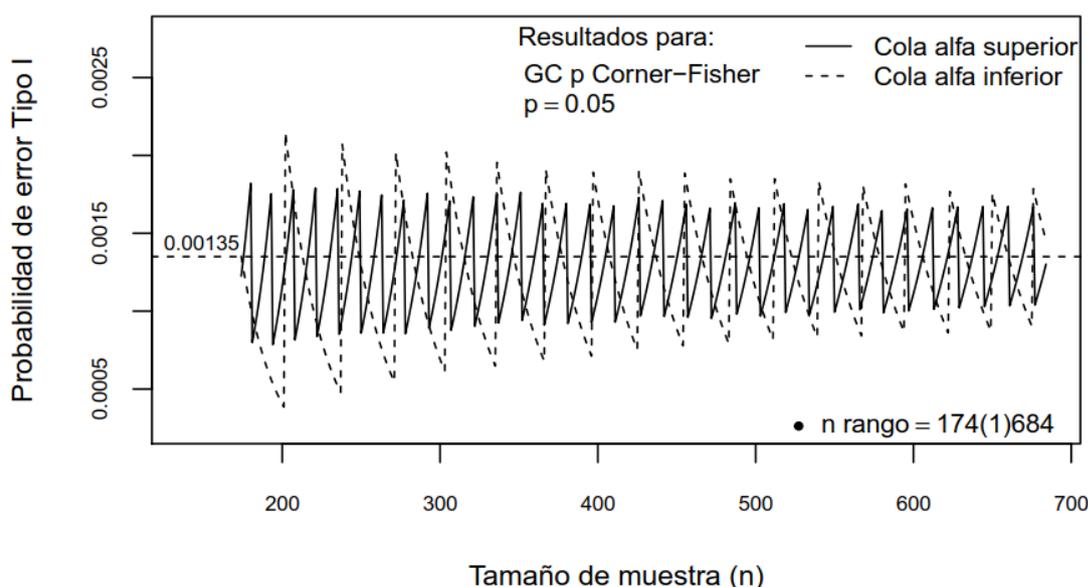


Figura 42. Ejemplo de colas alfa típicas para el GC p Corner-Fisher

La Figura 43 muestra el comportamiento típico del ARL_0 , como se puede apreciar para tamaños de muestra bajos el ARL_0 es, en muchos casos, excesivamente alto y de todas las posibles combinaciones, con $p = 0.05$ y n rango = 174(1)684, solo el 71% tendrán un ARL_0 aceptable (entre 250 y 450).

Los valores ARL_0 excesivamente altos se dan cuando las dos colas están muy por debajo de 0.00135, recordemos que $ARL_0 = 1/\alpha$ donde $\alpha = \alpha_I + \alpha_S$.

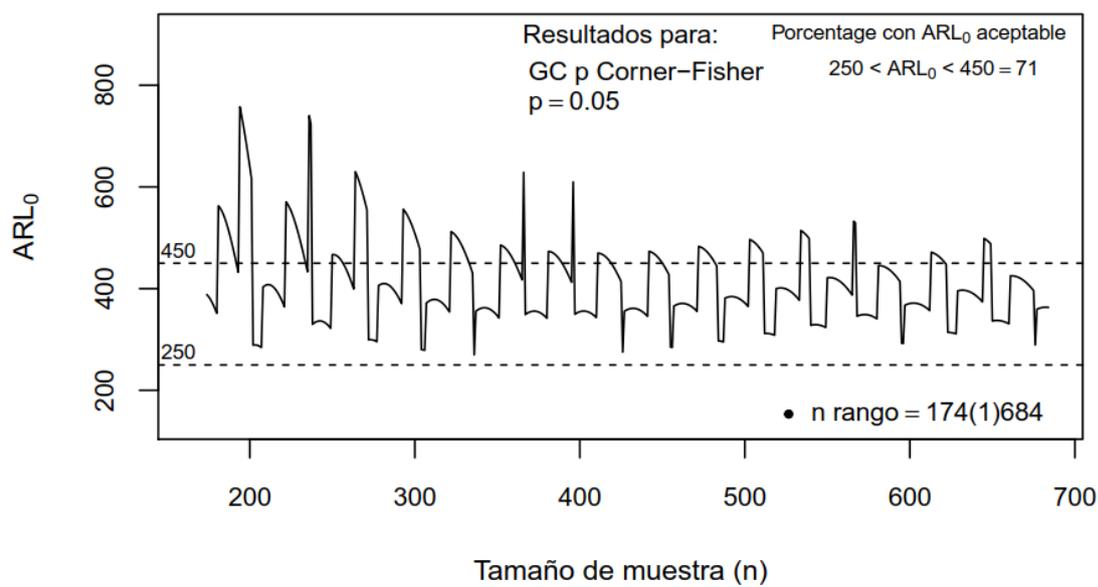


Figura 43. Ejemplo del comportamiento típico del ARL_0 para el GC p Corner-Fisher

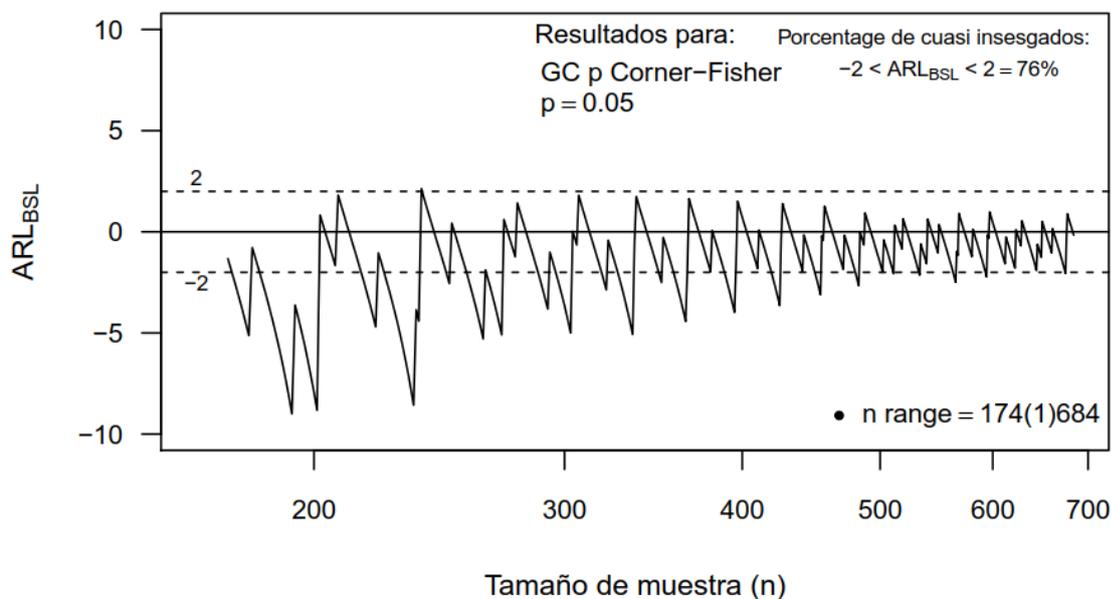


Figura 44. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Corner-Fisher

Por otro lado la Figura 44 muestra el comportamiento típico del ARL_{BSL} para el GC p Corner-Fisher, si se compara estos resultados con los obtenidos para el GC p Shewhart (ver Figura 40) es fácil deducir que existe una gran mejora. Sin embargo, los resultados también demuestran que del total de combinaciones, con $p = 0.05$ y n rango = 174(1)684, solo el 76% resultarían en gráficos de control con curvas ARL cuasi incesgadas y que los gráficos con curvas sesgadas se darían principalmente con tamaños de muestra pequeños.

La Tabla 13 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC p Corner-Fisher, antes de analizar esos resultados es necesario enfatizar que un gráfico de control es considerado como óptimo, en términos del ARL, solo cuando este tiene: i) una curva ARL cuasi incesgada y ii) un ARL_0 aceptable.

Los resultados demuestran que en promedio solo un 75% de gráficos construidos con el GC p Corner-Fisher tendrían curvas ARL cuasi incesgadas y que además, en promedio, solo un 78% de ellos tendrían valores ARL_0 aceptables. De esto se concluye que si se utilizase el GC p Corner-Fisher, habría un alto riesgo de trabajar con un gráfico que fuese no óptimo.

Tabla 13: Síntesis de los resultados obtenidos para el GC p Corner-Fisher

p	Rango de tamaños de muestra	Porcentaje de gráficos cuasi incesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi incesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
0.01	909(1)3264	75	80
0.02	450(1)1764	75	82
0.04	220(1)864	75	80
0.05	174(1)684	76	76
0.10	83(1)324	74	71
0.15	52(1)204	76	68

Cabe mencionar que un GC p basado en una pequeña variación al propuesto por Winterbottom fue presentado por Joekes&Pimentel-Barbosa (2013). Ellos utilizan los cuatro primeros términos de la expansión Corner-Fisher, a diferencia de Winterbottom que utiliza solo los tres primeros. Para determinar la diferencia que había entre estos dos gráficos, se realizó un estudio comparativo de sus ARL_{BSL} y ARL_0 . Los resultados demostraron que el GC de Winterbottom da mejores resultados que el propuesto por Joekes&Pimentel-Barbosa.

5.5.2.1.2 Gráfico de Control p R&S - basado en análisis de regresión

Ryan y Schwertman (1997) propusieron un GC p basado en una fórmula derivada por medio de análisis de regresión y que fue diseñada para computar límites de control equiparables a los límites 3-sigma del GC p Shewhart.

Los límites de control para el GC p R&S se calculan mediante la ecuación (28).

$$\text{Límites del GC } p \text{ R\&S} = \frac{1}{n} [a + b(np) + c\sqrt{np}] \quad (28)$$

Donde:

- Para el LCS usar: $a=0.6195$; $b=1.00523$; $c=2.983$
- Para el LCI usar: $a=2.9529$; $b=1.01956$; $c=3.2729$
- n = tamaño de muestra
- p = fracción promedio de no conformes del proceso
- Línea central del gráfico = p

Para caracterizar el comportamiento de las colas alfa para el GC p R&S se realizó un estudio igual al hecho para el GC p Corner-Fisher. La Figura 45 muestra un ejemplo de las colas alfa que se obtienen con los límites computados con (28), mientras que la Figura 46 muestra el comportamiento típico del ARL_0 . Como se puede apreciar para tamaños de muestra pequeños el ARL_0 es, en muchos casos, excesivamente bajo y de todas las posibles combinaciones, con $p = 0.05$ y n rango = 174(1)684, solo el 71% tendrían un ARL_0 aceptable (entre 250 y 450).

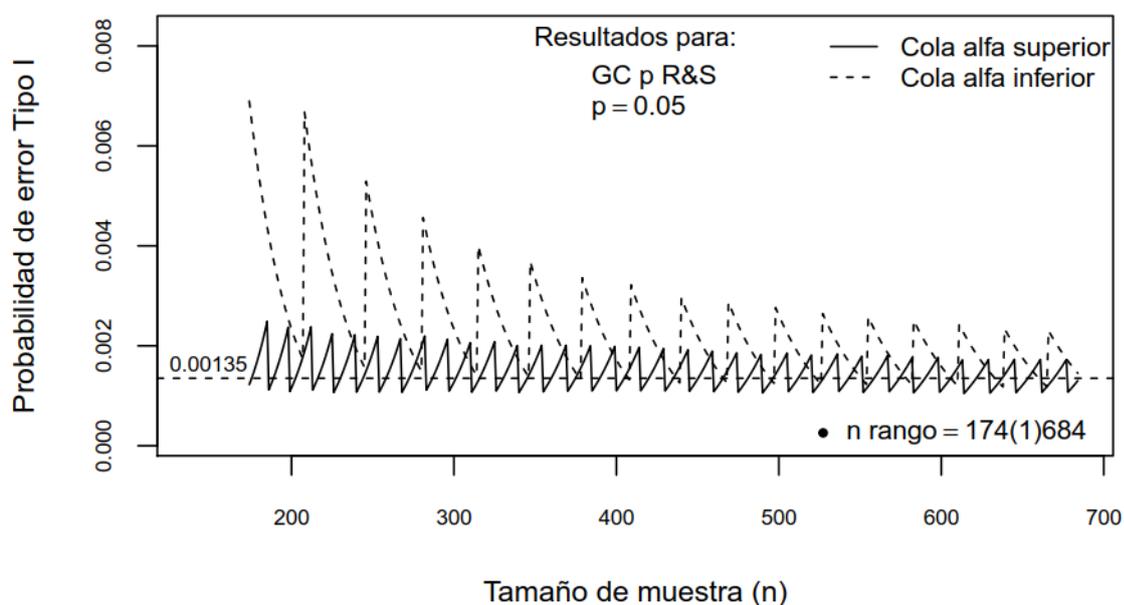


Figura 45. Ejemplo de colas alfa típicas para el GC p R&S

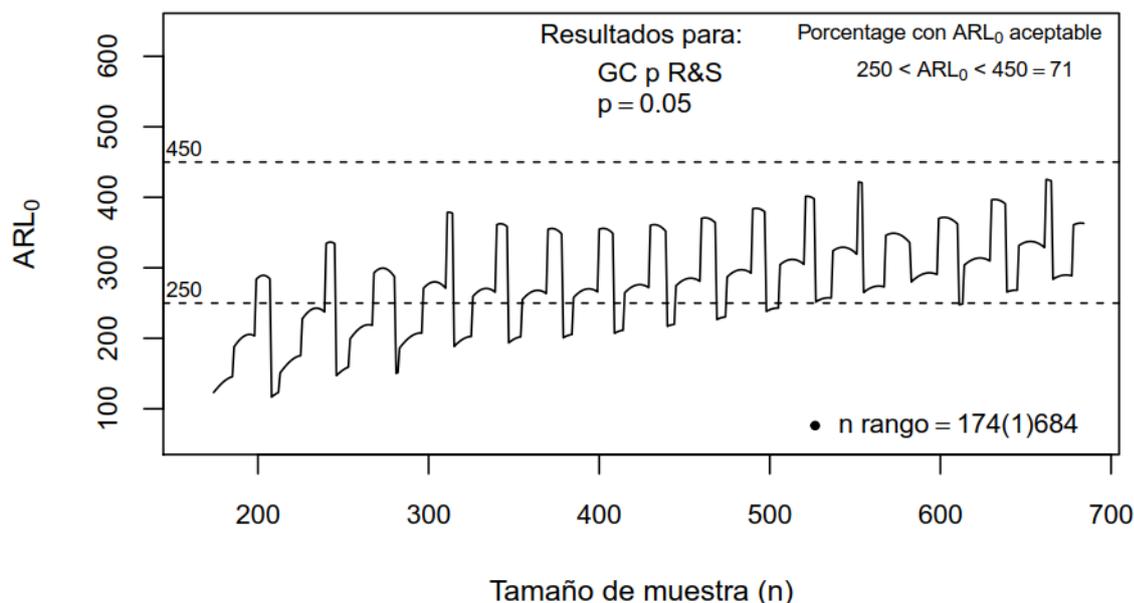


Figura 46. Ejemplo del comportamiento típico del ARL_0 para el GC p R&S

La Figura 47 muestra el comportamiento típico del ARL_{BSL} para el GC p R&S, si se compara estos resultados con los obtenidos para el GC p Shewhart (ver Figura 40) es muy claro ver que existe una gran mejora. Sin embargo, los resultados también demuestran que del total de combinaciones, con $p = 0.05$ y n rango = 174(1)684, solo el 84% darían gráficos de control con curvas ARL cuasi inesgadas.

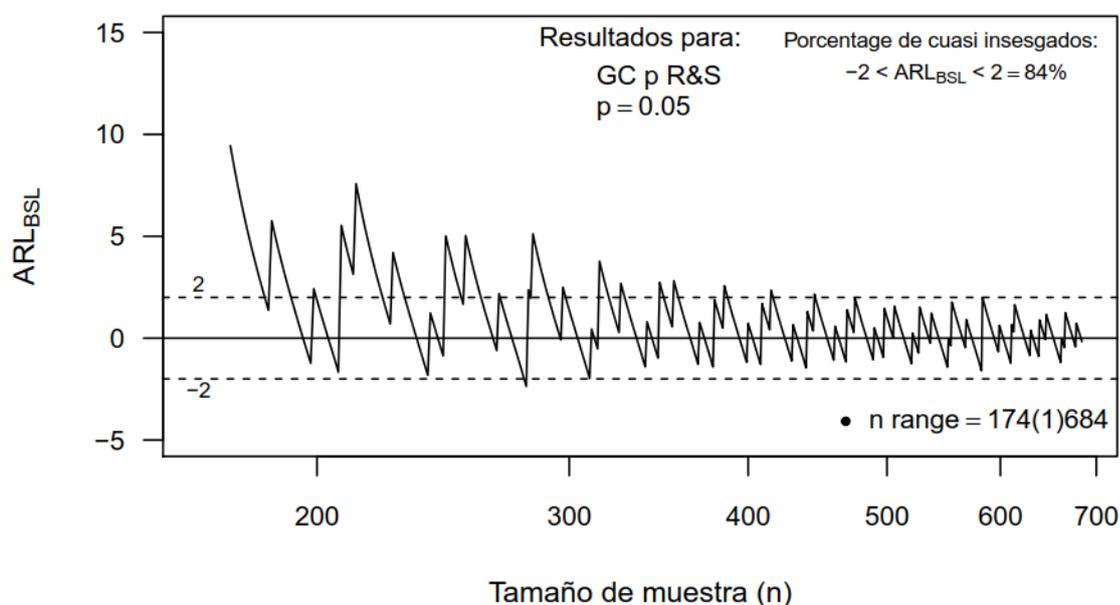


Figura 47. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p R&S

La Tabla 14 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC p R&S, como se mencionó anteriormente lo ideal es obtener gráficos de control con curvas ARL cuasi insesgadas y ARL_0 aceptables.

Los resultados demuestran que, para $p \leq 0.10$, una alta proporción de gráficos construidos con el GC p R&S tendrían curvas cuasi insesgadas, algo que en primera instancia parecería ser alentador. Sin embargo, los resultados del ARL_0 no son tan buenos, especialmente para valores de $p < 0.03$. De esto se concluye que si se utilizase el GC p R&S, habría un alto riesgo de trabajar con un gráfico que fuese no óptimo.

Tabla 14: Síntesis de los resultados obtenidos para el GC p R&S

p	Rango de tamaños de muestra	Porcentaje de gráficos cuasi insesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi insesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
0.01	909(1)3264	88	45
0.02	450(1)1764	87	55
0.04	220(1)864	85	79
0.05	174(1)684	84	81
0.10	83(1)324	76	81
0.15	52(1)204	63	20

5.5.2.2 Gráficos de control p : Enfoque uso de transformaciones no lineales

En la sección 4.1.2.3 se demostró que la distribución binomial es asimétrica cuando $p \ll 0.5$, y que en estos casos el ajuste de la aproximación normal a las colas de la distribución binomial es muy deficiente. Para tratar de superar este problema, y para crear GCs que tuviesen colas alfa más cercanas a 0.00135, se encontró que varios autores hicieron uso de la técnica llamada “transformación de datos”, mediante la cual datos no normales pueden ser aproximados a normales mediante una transformación. Se estableció que, para el caso de datos binomiales, tres tipos de transformaciones habían sido utilizadas, siendo estas: 1) La transformación arco seno 2) La transformación raíz cuadrada y 3) La transformación Q-binomial. Los GCs que pertenecen a este enfoque se detallan en la Tabla 15.

Tabla 15: Gráficos de control p alternativos: Enfoque uso de transformaciones

<i>Descripción del GC</i>	<i>Autores y publicación</i>
GC p basado en la transformación Arco Seno	Ryan (2011, pág. 188&189)
GC p basado en la transformación raíz cuadrada	Tsai, Lin y Wu (2006)
GC p basado en la transformación Q-binomial	Quesenberry (1991) Quesenberry (1995)

Los tres GCs incluidos en la Tabla 15 fueron sometidos a un estudio que tuvo como fin caracterizar el comportamiento de sus ARL_0 y ARL_{BSL} , y también determinar sus rendimientos en cuanto a la proporción de gráficos de control óptimos que se obtienen dentro de un rango de n . Los resultados de este estudio permitieron cumplir con el tercer objetivo específico de esta tesis.

Los resultados demostraron que el rendimiento ARL_{BSL} de estos GCs, en cuanto a curvas ARL cuasi insesgadas, era muy pobre. Recordemos que un gráfico de control puede ser considerado como óptimo, solo si tiene un ARL_0 aceptable y una curva ARL cuasi insesgada. En base a los resultados obtenidos se concluyó que los gráficos basados en transformaciones eran inferiores a los basados en optimización de límites. Por esta razón, y por motivos de concisión, se decidió presentar en esta tesis, aparte de los detalles de cada uno de estos GCs, solo una tabla en donde se resumen los resultados del ARL_{BSL} .

Cabe mencionar que los resultados de esta parte de la investigación fueron presentados en un congreso internacional, ver Argoti et al (2018b).

5.5.2.2.1 Gráfico de Control p basado en la transformación Arco Seno

Los detalles del GC p Arco Seno se dan en la ecuación (29). Como se puede apreciar, tanto los límites de control como el estadístico que se traza en el gráfico, se obtienen utilizando la función inversa del Seno.

$$\text{Límites del GC } p \text{ Arco Seno} = \text{Seno}^{-1}(\sqrt{p}) \pm \frac{3}{2\sqrt{n}} \quad (29)$$

$$\text{Estadístico trazado en el gráfico} = t_i = \text{Seno}^{-1}(\sqrt{y_i})$$

Donde:

- Para el LCS usar: +
- Para el LCI usar: -

- Línea central del gráfico = $\text{Seno}^{-1}(\sqrt{p})$
- $y_i = x_i/n_i \quad i = 1,2,3 \dots \rightarrow$ Fracción de no conformes por muestra
- $n =$ tamaño de muestra
- $p =$ fracción promedio de no conformes del proceso

Para poder computar las colas alfa y caracterizar el ARL_{BSL} y ARL_0 de este GC, se requiere primeramente que los límites del gráfico sean convencionales o no transformados, algo que se logra mediante la ecuación (30). Con los límites convencionales es posible obtener las colas alfa siguiendo el mismo procedimiento utilizado para el GC p Shewhart en la sección 4.1.6.1. Un ejemplo de los resultados obtenidos para las colas alfa se muestra en la Figura 48, como se puede apreciar la cola superior oscila por encima de 0.00135 mientras que la cola inferior oscila por debajo.

$$\text{Límites convencionales equivalentes} = \text{Sin}^2 \left[\text{Sin}^{-1}(\sqrt{p}) \pm \frac{3}{2\sqrt{n}} \right] \quad (30)$$

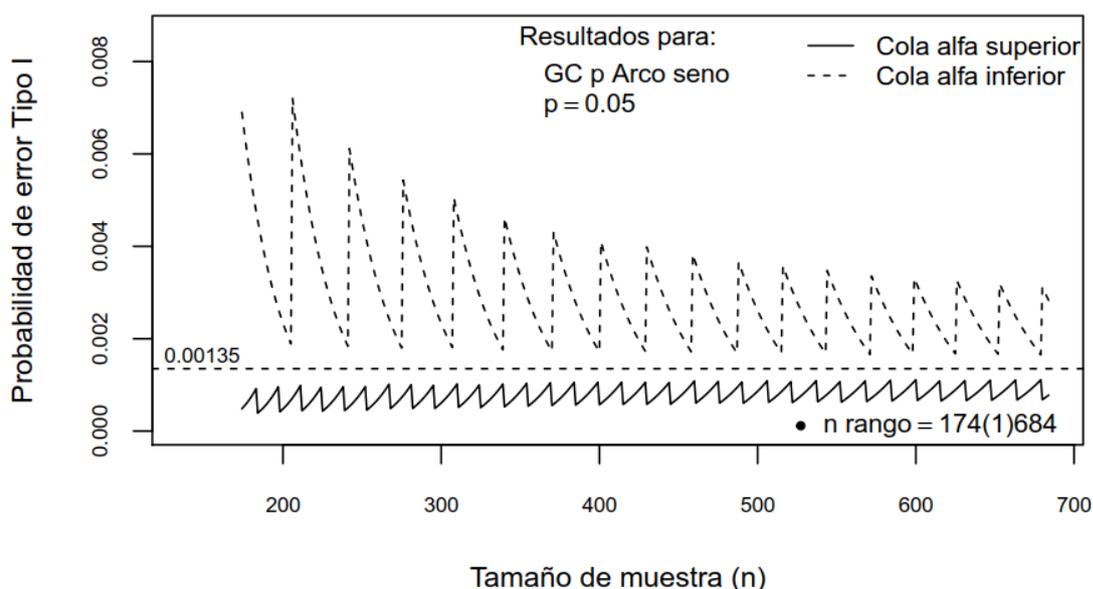


Figura 48. Ejemplo de colas alfa típicas para el GC p Arco Seno

Una vez obtenidas las colas alfa es posible computar el ARL_0 y el ARL_{BSL} . La Figura 49 muestra un ejemplo de los resultados del ARL_0 para $p = 0.05$, como se puede apreciar, para n pequeños, en muchas ocasiones el valor del ARL_0 es excesivamente bajo. Además, de todas las posibles combinaciones, para $p = 0.05$ y n rango = 174(1)684, solo el 69% tendrían un ARL_0 aceptable (entre 250 y 450).

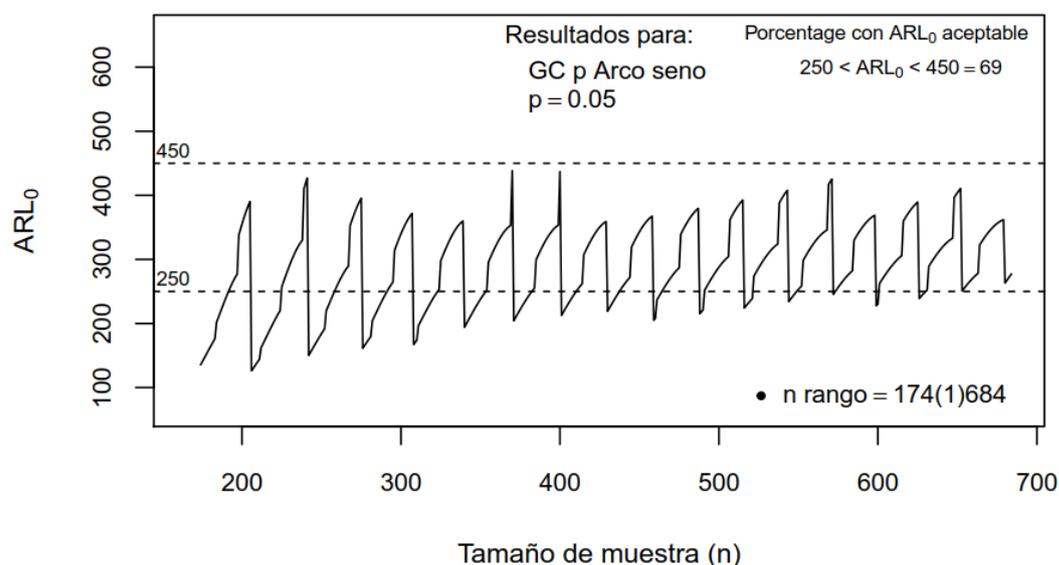


Figura 49. Ejemplo del comportamiento típico del ARL_0 para el GC p Arco Seno

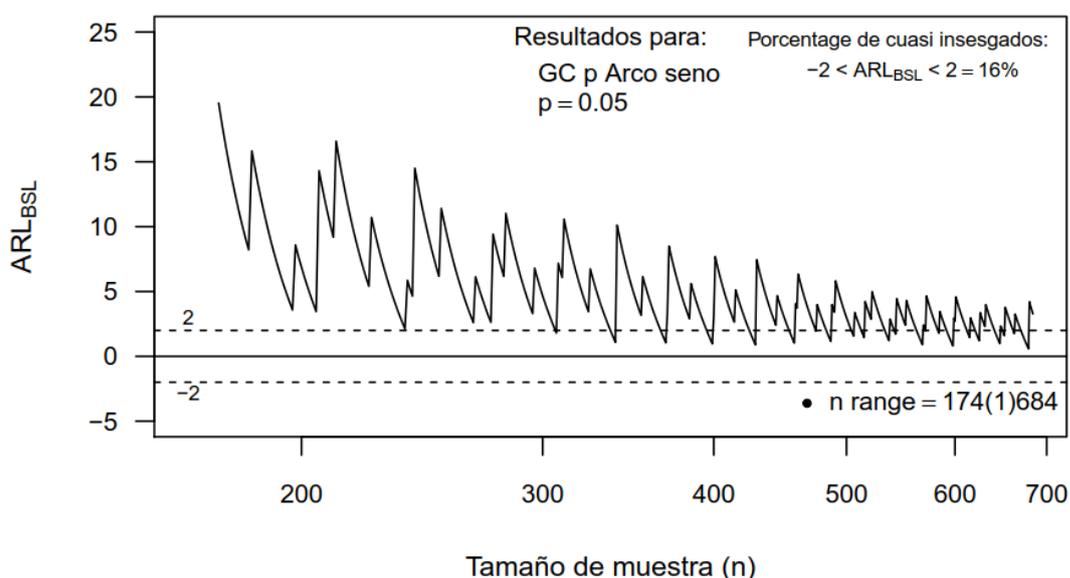


Figura 50. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Arco Seno

La Figura 50 muestra el comportamiento típico del ARL_{BSL} para el GC p Arco Seno, si se compara estos resultados con los obtenidos para el GC p Shewhart (ver Figura 40) es evidente que existe una mejora. Sin embargo, los resultados también demuestran que del total de combinaciones, con $p = 0.05$ y n rango = 174(1)684, solo el 16% serían gráficos de control con curvas ARL cuasi inesgadas, lo que representa un resultado muy pobre.

Los resultados del ARL_{BSL} también demuestran que la mayoría de gráficos construidos con el GC p Arco Seno tendrían curvas ARL con sesgo positivo, lo cual implica una menor capacidad para detectar deterioros en el proceso. Para ejemplificar este tipo de curvas se incluye la Figura 51, que muestra el caso de curvas ARL con sesgo positivo y negativo.

Por razones de concisión se presentan, en la sección 5.5.2.2.4, solo los resultados del rendimiento ARL_{BSL} para este GC. Sin embargo, cabe mencionar, que en la Tabla 18 se incluyen tanto los resultados del ARL_{BSL} como los del ARL_0 .

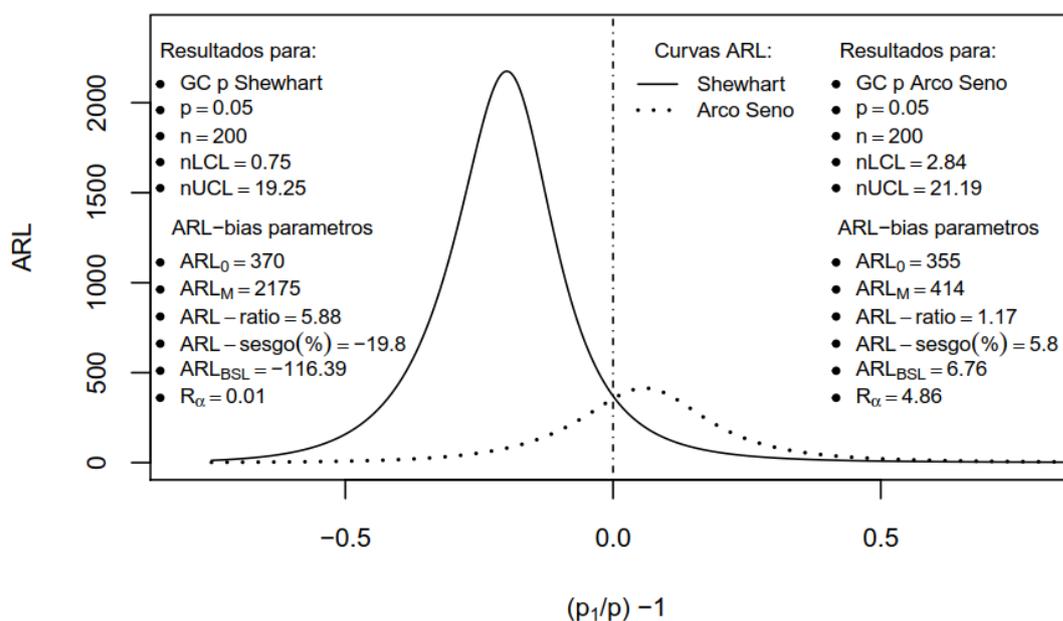


Figura 51. Ejemplo de curvas ARL con sesgo positivo y negativo

5.5.2.2.2 Gráfico de Control p basado en la transformación raíz cuadrada

Tsai, Lin y Wu (2006) propusieron un GC basado en la transformación raíz cuadrada (TRC), cuyos límites son computados mediante la ecuación (31). Para computar las colas alfa y caracterizar el ARL_{BSL} y ARL_0 de este GC, se requiere primeramente que los límites del gráfico sean convencionales o no transformados. Algo que se logra elevando al cuadrado los límites computados con la ecuación (31).

Con los límites convencionales se pueden obtener los valores de las colas alfa, tal como se hizo para el GC p Shewhart en la sección 4.1.6.1, y con estas es posible computar todos los parámetros de sesgo de la curva ARL.

Por razones de concisión se presentan, en la sección 5.5.2.2.4, solo los resultados del rendimiento ARL_{BSL} para este GC.

Límites del GC p TRC (31)

$$LCS = \sqrt{p} + \frac{3}{2} \sqrt{\frac{1-p}{n}} - \frac{1}{2} \left(\frac{1-p}{n\sqrt{p}} \right)$$

$$LCI = \sqrt{p} - \frac{3}{2} \sqrt{\frac{1-p}{n}} - \frac{9}{8} \left(\frac{1-p}{n\sqrt{p}} \right)$$

Donde:

- Línea central del gráfico = \sqrt{p}
- Estadístico trazado en el gráfico = $y_i = \sqrt{\frac{x_i}{n_i}}$
- *Limites convencionales equivalentes* = LCS^2 y LCL^2

5.5.2.2.3 Gráfico de control p basado en la transformación Q

Quesenberry (1991) propuso el GC p basado en la transformación Q , cuyos límites de control están fijados en ± 3 como se muestra en la ecuación (32), esto debido a que el estadístico que se traza sigue una distribución $N(0,1)$.

Para obtener las colas alfa para este GC hay que primeramente encontrar el número de no conformes (x) que corresponden a los límites ± 3 y luego utilizar estos para computar los valores alfa. Ejemplo de cómo hacer estos cálculos los da Quesenberry en sus publicaciones (1991, pág. 6) y (1995, pág. 209).

Por razones de concisión se presentan, en la sección 5.5.2.2.4, solo los resultados del rendimiento ARL_{BSL} para este GC.

Limites del GC p Q (32)

$$LCS = +3$$

$$LCL = -3$$

Donde:

- Línea central del gráfico = 0
- Estadístico trazado en el gráfico = Q_i
- $Q_i = \Phi^{-1}(u_i)$
- Φ^{-1} = la inversa de la distribución normal estándar
- $u_i = b(x_i, n_i, p)$ for $i = 1, 2, 3 \dots$

5.5.2.2.4 Comparativa de los resultados ARL_{BSL} para los gráficos p basados en transformaciones

En esta sección se presenta la Tabla 16, en donde se resumen los resultados obtenidos para la proporción de gráficos con curvas ARL “cuasi insesgadas” que se alcanzan con los GCs basados en transformaciones. Cabe mencionar que para identificar si una curva es o no cuasi insesgada, se utilizó el criterio $-2 < ARL_{BSL} < 2$ dado en la sección 5.4.2.

La Tabla 16 muestra solo los resultados para $p = 0.01, 0.05, 0.1$ y 0.15 , estos resultados son suficientes para ejemplificar el rendimiento que se obtiene con los GCs p basados en transformaciones. Como se puede apreciar el porcentaje de gráficos cuasi insesgados que se

obtienen con estos GCs es extremadamente bajo. De esto es posible concluir que si se empleara cualquiera de estos GCs, habría un alto riesgo de que el GC utilizado fuese no óptimo.

Además, también es posible concluir que, debido a su pobre rendimiento ARL_{BSL} , estos gráficos son inferiores a los basados en optimización de límites (ver sección 5.5.2.1).

Tabla 16: Resultados ARL_{BSL} para GCs p basados en transformaciones

<i>GC</i>	<i>p</i>	Rango de tamaños de muestra	Porcentaje de gráficos con curvas ARL cuasi insesgadas $-2 < ARL_{BSL} < 2$
Arcoseno	0.01	978(1)3564	12%
TRC			15%
Q			18%
Arcoseno	0.05	189(1)684	17%
TRC			22%
Q			16%
Arcoseno	0.1	90(1)324	21%
TRC			33%
Q			15%
Arcoseno	0.15	57(1)204	28%
TRC			45%
Q			13%

5.5.2.3 Sinopsis de la sección 5.5.2

Los resultados obtenidos en la fase de investigación resumida en la sección 5.5.2, dejan en evidencia que existen varios GC p alternativos que son superiores al GC p Shewhart. Sin embargo, una conclusión muy importante derivada de los resultados, es que estos GCs también pueden tener curvas ARL sesgadas y/o ARL_0 no aceptables. De esto se deduce que si se usara cualquiera de estos GCs alternativos, se correría el riesgo de trabajar con un gráfico de control no óptimo.

5.5.3 El gráfico de control p K_{mod}

Esta sección resume el trabajo realizado para cumplir con el cuarto objetivo específico de esta tesis que fue “Desarrollar nuevas propuestas de gráficos de control por atributos”. Para lograr este objetivo se creó el GC p K_{mod} , que sirve para monitorizar procesos a partir de la distribución binomial. Este GC tiene un rendimiento, en cuanto a la proporción de gráficos idóneos, superior a todos los GCs p alternativos analizados anteriormente y además cuenta con un método muy sencillo para verificar si sus curvas ARL son o no “cuasi insesgadas”.

Aparte de lo que se presenta en esta sección acerca del GC p K_{mod} , en la publicación de Argoti&Carrión-García (2019a), se dan otros detalles que podrían ser de interés.

5.5.3.1 El factor K adaptativo y los límites de control para el gráfico de control p K_{mod}

En la sección 5.5.1.1 se concluyó que para obtener valores ARL_{BSL} cercanos a cero es necesario reducir, a lo mínimo posible, la disparidad entre las colas alfa. Basados en este antecedente, y luego de una búsqueda y estudio de técnicas novedosas para reducir esa disparidad, se determinó que los mejores resultados se obtenían añadiendo el término $(\lambda_{I|S})/\sqrt{np(1-p)}$ al factor $K = 3$ de la ecuación para el GC p Shewhart mostrada en la Tabla 3 de la sección 4.1.4.1. Nótese que el término que se añade a $K = 3$ está basado en la desviación típica de la distribución binomial.

El factor K modificado resultante es $[3 \pm (\lambda_{I|S})/\sqrt{np(1-p)}]$, donde los signos más y menos están relacionados con los límites de control inferior y superior respectivamente, y $\lambda_{I|S}$ es una constante que debe ser determinada para cada límite de control (λ_I para el límite inferior y λ_S para el superior). A este factor K modificado se lo ha denominado K_{mod} , ver ecuación (33).

Para establecer valores óptimos de λ_I y λ_S se realizó un exhaustivo estudio de optimización heurística basado en los siguientes criterios:

1. Ratio de las colas alfa (R_α) lo más cercanos a 1
2. ARL_{BSL} lo más cercano a cero
3. ARL_0 entre 250 y 450

El estudio consistió en primeramente dar valores entre 0.6(0.1)2 a cada una de las constantes λ_I y λ_S de la ecuación (33). Luego de eso, y para cada uno de esos valores, se computaron individualmente las colas alfa (inferior y superior) que se obtenían para $p = 0.01(0.01)0.2$ usando n rangos específicos para cada p , y cuyos valores mínimo y máximo fueron

determinados siguiendo los mismos lineamientos utilizados para el análisis de las colas alfa del GC Shewhart (ver sección 4.1.6.1). Una vez obtenidos los resultados de las colas alfa, se identificaron los valores más destacados de λ_I y λ_S , es decir aquellos con los que, en promedio y para cada p , se obtenían colas alfa lo más cercanas a 0.00135. Una vez identificados los valores λ_I y λ_S destacados, se procedió a combinarlos metódicamente en (33) para luego, y con cada una de estas combinaciones, computar los R_α , ARL_{BSL} y ARL_0 para los mismos valores de p y n mencionados anteriormente. Una vez hecho esto, y con base en los criterios de optimización descritos previamente, se analizaron los resultados hasta dar con los valores óptimos de λ_I y λ_S .

Luego de un extenso análisis de los resultados mencionados al final del párrafo anterior, se determinó que con los valores $\lambda_I = 1$ para el Kmod del LCS y $\lambda_S = 1.6$ para el Kmod del LCI, se obtenían los mejores resultados. La ecuación (33) sirve para computar los límites de control del nuevo gráfico de control al que se lo ha denominado como p Kmod.

$$GC\ p\ Kmod\ Limites\ de\ control = p \pm (Kmod_{LCI|LCS}) \sqrt{\frac{p(1-p)}{n}} \quad (33)$$

Dónde:

Para el limite de control superior (LCS):

$$Kmod_{LCS} = 3 + \frac{\lambda_I}{\sqrt{np(1-p)}} ; \lambda_I = 1$$

Para el limite de control inferior (LCI):

$$Kmod_{LCI} = 3 - \frac{\lambda_S}{\sqrt{np(1-p)}} ; \lambda_S = 1.6$$

Cabe mencionar que el GC p Kmod ha sido optimizado para que sus mejores resultados se den dentro del rango $0.005 \leq p \leq 0.10$, esto debido a que se consideró que ese rango de p es mucho más común en procesos industriales modernos que $p > 0.10$. El rendimiento ARL (ARL_0 y ARL_{BSL}) del GC p Kmod comienza a disminuir paulatinamente a partir de $p = 0.10$, razón por la cual se recomienda que este GC no sea usado para valores de p por encima de 0.20.

Para ejemplificar y comparar los valores de las colas alfa que se obtienen con el GC p Kmod, se utilizan los ejemplos dados para el GC p Shewhart en la Figura 22 y en el Anexo 2. En el ejemplo de la Figura 22 se muestra que para $n = 100$ y $p = 0.1$, el GC Shewhart tiene límites iguales a $nLCL=1$ y $nUCL=19$ con los que se obtienen colas alfa iguales a $\alpha_I = 0.00032$ y $\alpha_S = 0.00198$, que resultan en un $R_\alpha = 0.16$. Ahora bien, para este caso el GC Kmod computa límites iguales a $nLCL = 2.6$ y $nUCL = 20$, con los que se obtienen $\alpha_I = 0.00194$, $\alpha_S = 0.00080$

y un $R_\alpha = 2.4$. Observando la Figura 22 se puede apreciar que con los límites Kmod las colas alfa toman los valores más cercanos posibles a 0.00135, lo que resulta en colas alfa con una menor disparidad en comparación a las obtenidas con el GC Shewhart.

Por otra parte, en el ejemplo expuesto en el Anexo 2 se muestra que, para $n = 600$ y $p = 0.05$, el GC Shewhart tiene límites iguales a $nLCL=13.98$ y $nUCL= 46.02$, con los que se obtienen $\alpha_I = 0.00031$, $\alpha_S= 0.00189$ y $R_\alpha = 0.166$. Para este caso el GC Kmod computa límites iguales a $nLCL = 15.58$ y $nUCL= 47$, con los que se obtienen $\alpha_I= 0.00158$, $\alpha_S= 0.0011$ y $R_\alpha = 1.4$. A través de las figuras incluidas en el Anexo 2 es posible deducir que con los límites Kmod, las colas alfa toman los valores más cercanos posibles a 0.00135 lo que, como en el ejemplo anterior, resulta en colas alfa con una menor disparidad en comparación a las obtenidas con el GC Shewhart.

5.5.3.2 Gráfico de control p Kmod: Caracterización del ARL_0 y ARL_{BSL}

El GC p Kmod fue sometido a un estudio que tuvo como fin caracterizar el comportamiento del ARL_0 y ARL_{BSL} , y también determinar su rendimiento en cuanto a la proporción de gráficos de control óptimos que se obtienen dentro de un rango de n . Para realizar este estudio se utilizaron los mismos lineamientos, para el rango de p y n , que los empleados en la sección 5.5.1 para el GC p Shewhart.

La Figura 52 muestra un ejemplo de las colas alfa típicas que se obtienen con los límites computados con (33). Comparando estas colas alfa con las del GC p Shewhart, ver Anexo 1 caso $p = 0.05$, es fácil darse cuenta que las colas alfa del GC p Kmod están más cerca de 0.00135.

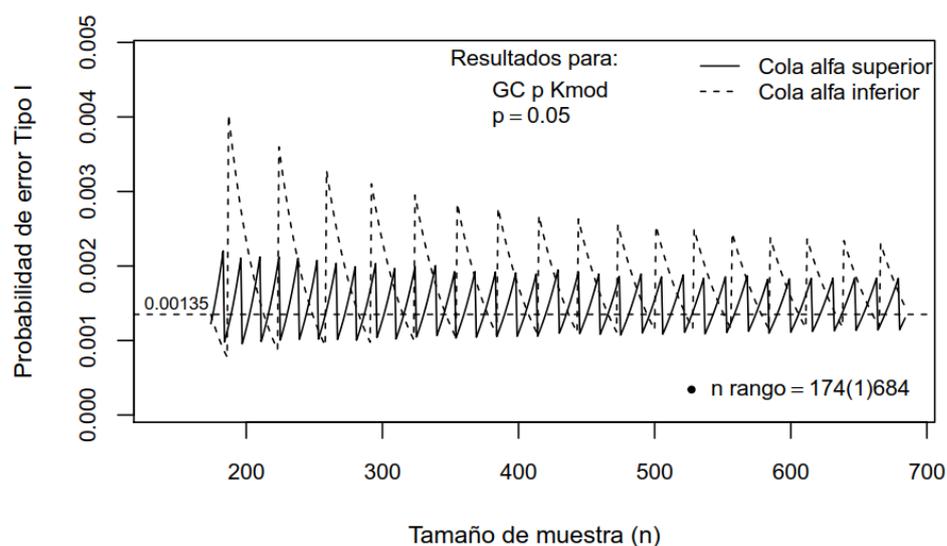


Figura 52. Ejemplo de colas alfa típicas para el GC p Kmod

La Figura 53 muestra el comportamiento del ARL_0 para $p = 0.05$, como se puede observar 88% de los graficos de control contruidos con $p = 0.05$ y n rango = 174(1)684 tendrían valores ARL_0 aceptables. Por otra parte la Figura 54 muestra que 84% de esos graficos tendrían curvas ARL cuasi insesgadas.

Resultados del ARL_0 y ARL_{BSL} para $p = 0.01$ y 0.10 se incluyen en el Anexo 6.

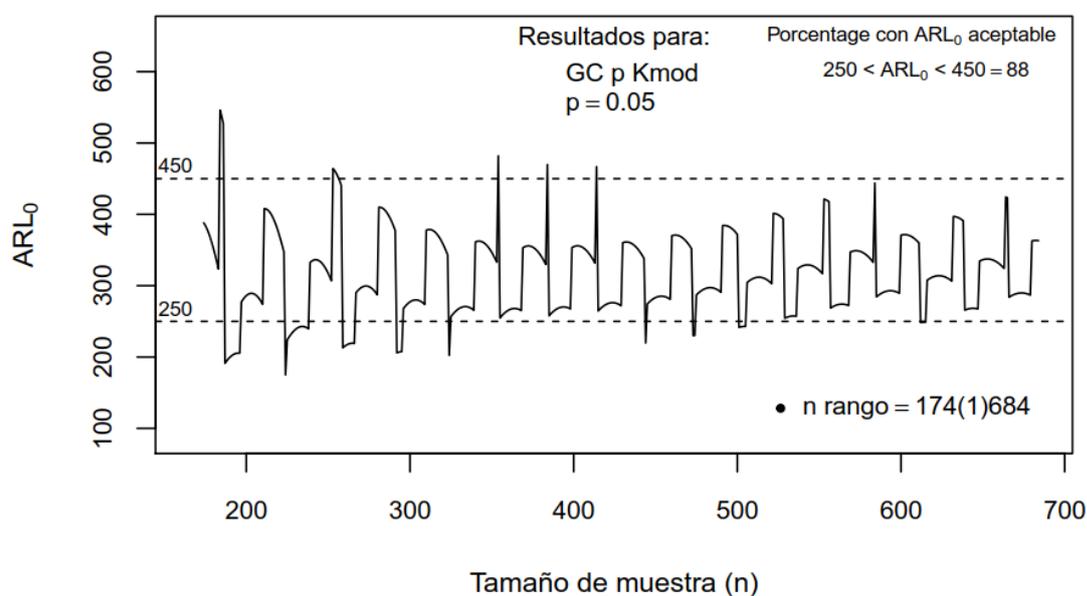


Figura 53. Ejemplo del comportamiento típico del ARL_0 para el GC p Kmod

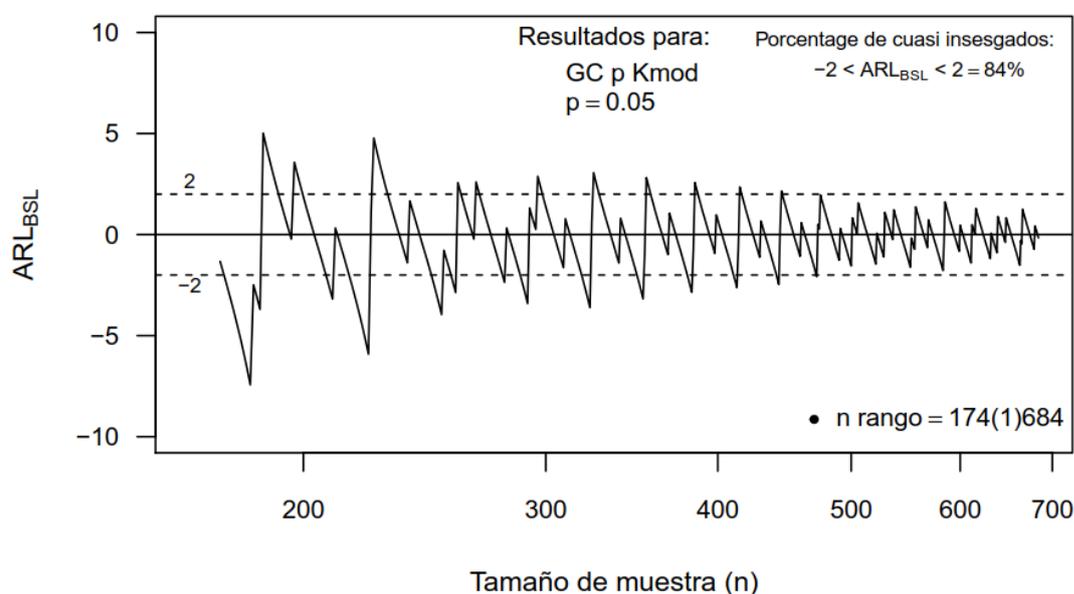


Figura 54. Ejemplo de comportamiento típico del ARL_{BSL} para el GC p Kmod

La Tabla 17 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC p Kmod. Cabe recalcar que un gráfico de control es

considerado como óptimo, solo cuando tiene una curva ARL cuasi incesgada y un ARL_0 aceptable.

Los resultados muestran que, para $p \leq 0.10$, en promedio un 83% de gráficos construidos con el GC p Kmod tendrían curvas ARL cuasi incesgadas, y que además, en promedio, un 89% de ellos tendrían valores ARL_0 aceptables. Estas proporciones en términos generales son más altas que las de cualquier otro GC p alternativo incluido en esta tesis.

Tabla 17: Síntesis de los resultados para el GC p Kmod

p	Rango de tamaños de muestra	Porcentaje de gráficos cuasi incesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi incesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
0.01	909(1)3264	82	87
0.02	450(1)1764	84	87
0.04	220(1)864	83	93
0.05	174(1)684	84	90
0.10	83(1)324	82	88
0.15	52(1)204	63	82

5.5.3.3 Gráfico de control p Kmod: Comparación con otros gráficos de control

La Tabla 18 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC p Kmod y con los GCs alternativos más destacados, ver sección 5.5.2. Antes de comentar acerca de los resultados, cabe recordar que en esta tesis un gráfico de control es considerado como óptimo solo cuando este tiene: i) una curva ARL cuasi incesgada y ii) un ARL_0 aceptable.

Si se analiza y compara los resultados de cada GC en cada p de la Tabla 18, es fácil darse cuenta que con el GC p Kmod se obtienen las proporciones más altas de gráficos cuasi incesgados con ARL_0 adecuados. Por ejemplo, para $p = 0.05$ el porcentaje de cuasi incesgados es 83% y de estos un 95% tendrán un ARL_0 aceptable; ningún otro GC sobrepasa estos porcentajes. Por lo que, basado en los resultados expuestos en esa tabla, es posible concluir que el GC p Kmod es superior a cualquiera de los GCs p incluidos en esta tesis.

Sin embargo, de los resultados también es posible concluir que con el GC p Kmod, todavía existe el riesgo, aunque mucho menor que con los otros GCs, de obtener gráficos de control con curvas ARL sesgadas y/o ARL_0 inaceptables. Ahora bien, computar el ARL_0 para

verificar si es aceptable o no es relativamente sencillo, ya que solo se necesita computar los valores de las colas alfa. Sin embargo, obtener el ARL_{BSL} es mucho más complejo ya que se requiere computar los parámetros de la curva ARL. Con esto en mente se concibió un método que permite determinar de una manera muy fácil, si un GC p Kmod tiene o no una curva ARL insesgada. Este método se detalla en la siguiente sección.

Tabla 18. Síntesis de los resultados para el GC p Kmod y para los GCs p alternativos más destacados

GC	p	n rango min max	% de gráficos con curvas ARL dentro de $-2 < ARL_{BSL} < 2$	Rendimiento del ARL_0 para gráficos que tienen curvas ARL dentro de $-2 < ARL_{BSL} < 2$					
				ARL_0 cuartiles (%)					% dentro de $250 < ARL_0 < 450$
				0	25	50	75	100	
Kmod	0.01	909 3564	82%	173	286	306	353	463	87%
R&S			88%	123	222	242	286	363	45%
CF			75%	276	344	384	432	605	80%
Arco seno			11%	284	332	358	377	416	100%
Kmod	0.02	450 1764	84%	180	272	311	352	484	87%
R&S			87%	128	232	254	299	381	55%
CF			75%	251	343	393	429	581	82%
Arco seno			11%	277	327	349	371	412	100%
Kmod	0.04	220 864	83%	195	272	315	346	480	93%
RB			85%	139	255	278	331	404	79%
CF			75%	266	344	381	444	601	80%
Arco seno			14%	285	331	345	368	417	100%
Kmod	0.05	174 684	84%	175	274	309	356	482	90%
R&S			84%	145	267	290	345	425	81%
CF			76%	270	352	378	445	630	76%
Arc seno			16%	276	336	355	367	425	100%
Kmod	0.10	83 324	82%	181	269	315	352	531	88%
R&S			76%	219	344	379	456	555	81%
CF			74%	259	348	381	457	623	71%
Arc seno			20%	261	334	346	372	428	100%
Kmod	0.15	52 204	63%	211	269	324	342	490	82%
R&S			63%	331	457	500	614	753	20%
CF			76%	282	339	387	461	690	68%
Arc seno			27%	268	329	345	367	430	100%

5.5.3.4 Gráfico de control p Kmod: Método para determinar si una curva ARL es cuasi insesgada

En esta sección se expone un método que permite verificar, sin la necesidad de computar los parámetros de la curva ARL, si un gráfico GC p Kmod tiene o no una curva ARL cuasi insesgada. Para ilustrar como funciona este método, se incluyen la Figura 55 y la Figura 56, que muestran la relación que existe entre el ARL_{BSL} y el ratio de las colas alfa (R_α) para $p = 0.05$ y $p = 0.10$

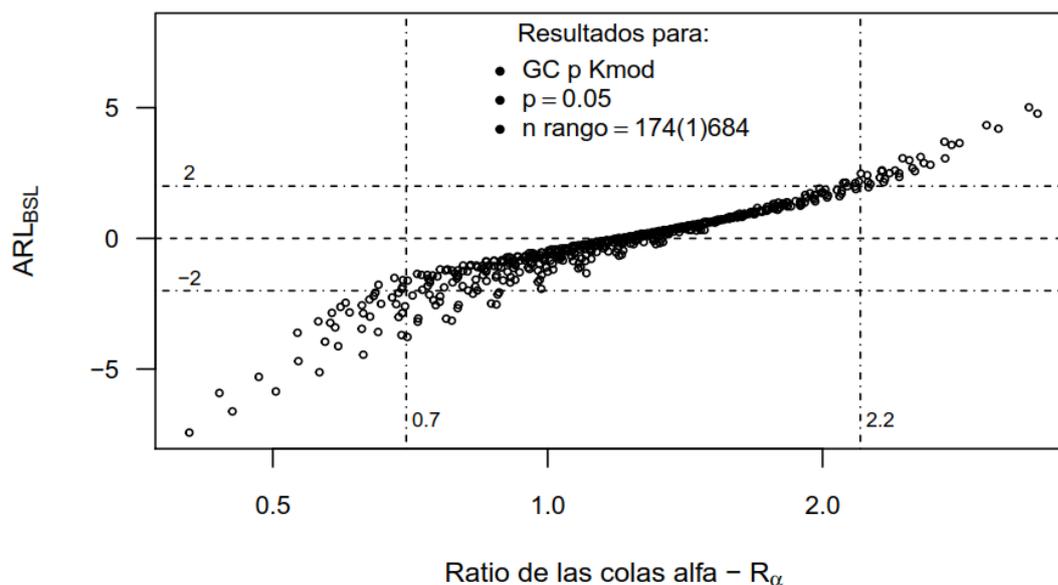


Figura 55. GC p Kmod: Relación entre el ARL_{BSL} y R_α para $p = 0.05$

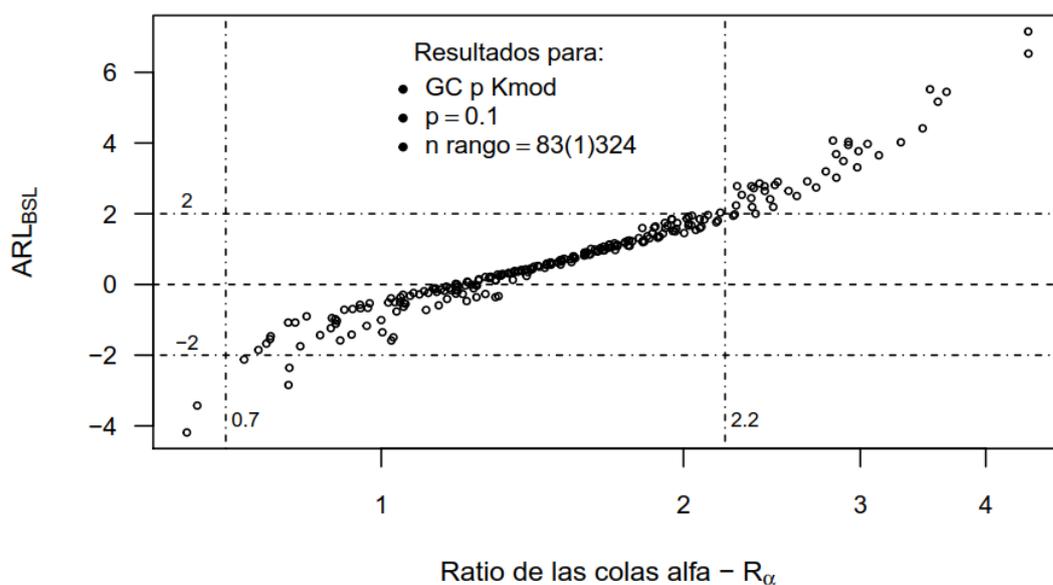


Figura 56. GC p Kmod: Relación entre el ARL_{BSL} y R_α para $p = 0.10$

En esas figuras se han trazado los límites $R_\alpha = 0.7$ y $R_\alpha = 2.2$ dentro de los cuales, y como se puede observar, la gran mayoría de combinaciones de p y n están dentro del criterio $-2 < ARL_{BSL} < 2$ para curvas cuasi ARL insesgadas. Estos límites fueron obtenidos luego de un extenso análisis de resultados de ARL_{BSL} vs R_α entre $p = 0.005$ y $p = 0.20$. El método para determinar si un GC p Kmod tiene una curva ARL cuasi insesgada o no, consiste simplemente en verificar si el valor del ratio de sus colas alfa cae dentro de los límites $0.7 < R_\alpha < 2.2$.

Sin embargo, cabe mencionar que en ocasiones aun cuando el R_α está entre 0.7 y 2.2 , el ARL_{BSL} para ciertas combinaciones de p y n podría exceder los límites $ARL_{BSL} = \pm 2$. En los análisis realizados se identificó que por lo general cuando esto sucede, los límites excedidos no sobrepasan los valores $ARL_{BSL} = \pm 3$, por lo que el sesgo de las curvas ARL para estos casos, podría considerarse que está dentro en un límite tope de aceptabilidad.

Una recomendación que se da es que cuando se obtengan valores de R_α muy cercanos a 0.7 , se determine el ARL_{BSL} , esto debido a que en ocasiones en ese límite pueden obtenerse valores ARL_{BSL} muy por debajo de -2 .

Cabe mencionar que el método descrito en esta sección, se enmarca dentro de lo requerido para cumplir con el segundo objetivo específico de esta tesis.

5.5.3.5 Gráfico de control p Kmod: ejemplo de curvas ARL

La Figura 57 muestra las curvas ARL obtenidas con el GC p Kmod para $p = 0.05$ con $n = 220$ y $n = 600$. Estos valores de p y n son los mismos que los utilizados para obtener las curvas ARL de la Figura 24 con el GC p Shewhart (ver sección 4.1.8.2). Comparando las curvas ARL de esas dos figuras (57 y 24), es fácil deducir que con el GC p Kmod la severidad del sesgo se ha reducido considerablemente.

En la Figura 57 se puede observar que la curva perteneciente a $n = 600$ tiene un $ARL_{BSL} = 0.4$ y un $ARL_0 = 372$, que son valores que denotan una curva ARL cuasi insesgada con un ARL_0 aceptable. Estos valores permiten deducir que el gráfico de control al cual pertenecen es óptimo, o en otras palabras, que puede detectar con la misma capacidad tanto mejoras como deterioros del proceso.

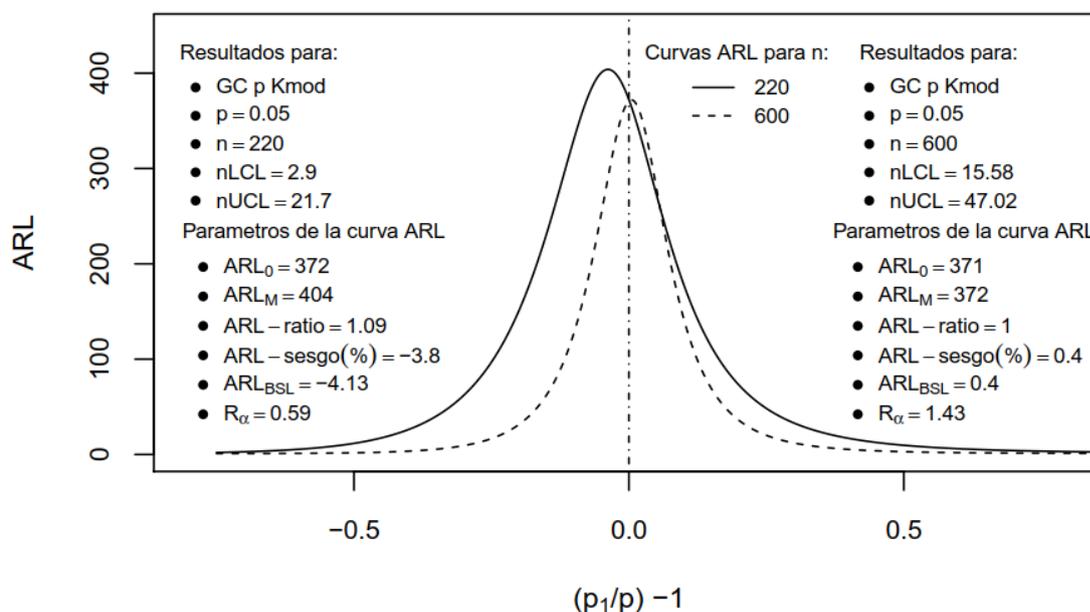


Figura 57. GC p Kmod: Ejemplo de curvas ARL para $p = 0.05$ con $n = 220$ y 600

Por otra parte, en la misma Figura 57, se puede apreciar que la curva perteneciente a $n = 220$ tiene un $ARL_{BSL} = -4.13$, lo que indica que tiene un sesgo inaceptable. Esta curva sirve de ejemplo para demostrar lo siguiente:

i. El uso del ratio de las colas alfa:

Para esta curva el $R_\alpha = 0.59$, lo que está muy por debajo del límite $R_\alpha = 0.7$ del método descrito en la sección 5.5.3.4. Es decir, con tan solo computar el R_α se hubiese podido identificar correctamente que el sesgo de la curva para $n = 220$ y $p = 0.05$ sería inaceptable.

ii. Cambiar el tamaño de muestra para obtener ARL_{BSL} y/o ARL_0 aceptables:

Una recomendación que se da en esta tesis en caso de que se obtengan GCs p Kmod con valores ARL_{BSL} y/o ARL_0 inaceptables, es que se modifique el tamaño de muestra hasta que estos se vuelvan aceptables.

Por ejemplo, para dar solución al caso de $n = 220$ y $p = 0.05$, se modificó el tamaño de muestra hasta dar con $n = 215$, con el que se obtienen $R_\alpha \approx 1$ y $ARL_0 \approx 400$. Es decir, con tan solo reducir el tamaño de muestra de 220 a 215, se obtendría un gráfico p Kmod con valores de sesgo aceptables, ver Figura 58.

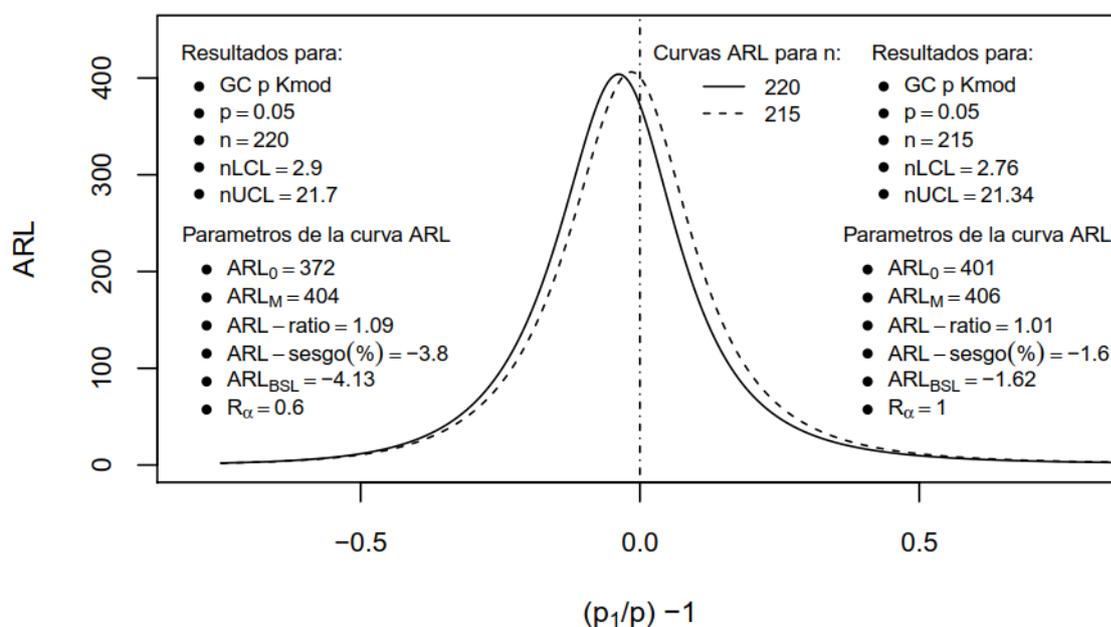


Figura 58: Kmod p chart: Curvas ARL para $p = 0.05$ con $n = 220$ y 215

5.5.3.6 Gráfico de control p Kmod: Aplicación al caso de monitorización de procesos mediante tamaños de muestra variables

En muchas ocasiones es preferible, o más conveniente, monitorizar un proceso mediante tamaños de muestras variables ($n_i : i = 1, 2, 3, \dots$). Para este caso, y asumiendo que p es constante, el procedimiento estándar es establecer límites de control para cada n_i , lo que resulta en un gráfico con límites de control variables, ver Montgomery (2009, pág. 301).

Ahora bien, se conoce que la distribución binomial adquiere valores discretos específicos para cada combinación de p y n . Por lo que, si se utilizara tamaños de muestra variables, el valor de las colas alfa variaría y sería específico para cada n_i . Lo que en consecuencia resultaría en valores de ARL_0 y ARL_{BSL} también específicos para cada tamaño de muestra usado.

Nuestra recomendación para este caso, es que se utilice el GC p Kmod de tal forma que con los n_i usados se obtengan siempre curvas ARL cuasi insesgadas y ARL_0 aceptables. Esto equivaldría a monitorizar el proceso con una curva ARL cuasi insesgada, cuya capacidad de detección variaría acorde a n_i .

Algo que se debe tomar en cuenta, es que la sensibilidad del gráfico para detectar cambios en el proceso tiene una relación directa con el tamaño de muestra; mientras más grande sea el tamaño de muestra, mayor será esa sensibilidad.

5.6 El gráfico de control u Kmod: Un nuevo gráfico para monitorizar procesos con base en la distribución de Poisson

La sección 5.6 agrupa todos los resultados relevantes que forma parte de un manuscrito, que al momento, está en revisión en una revista científica, ver Argoti&Carrión-García (2019b).

5.6.1 Gráfico de control u Shewhart: Caracterización del ARL_{BSL}

En esta sección se resumen los resultados del estudio realizado para caracterizar el comportamiento del ARL_{BSL} para el GC u Shewhart. Los antecedentes del por qué se realizó este estudio se exponen en la sección 4.2.6.

En la sección 5.4 se definió el parámetro ARL_{BSL} que sirve para cuantificar la severidad del sesgo de cualquier curva ARL, además ahí también consta la Tabla 11 que contiene los criterios para su interpretación. El ARL_{BSL} fue utilizado para analizar la severidad del sesgo de las curvas ARL de todos los GCs u incluidos en esta tesis, además también se utilizó el criterio $-2 < ARL_{BSL} < 2$, dado en la sección 5.4.2, para identificar curvas ARL *cuasi insesgadas*, o en otras palabras, curvas ARL cuyo sesgo tiene un impacto insignificante en la capacidad de monitorización de los gráficos de control a los que pertenecen.

Ejemplos de curvas ARL provenientes de gráficos de control obtenidos con el GC u Shewhart se dan en la Figura 31 y en las figuras incluidas en el Anexo 8. Como se puede observar, los valores ARL_{BSL} de esas curvas están muy por debajo de -2 , lo que lleva a la conclusión de que esas curvas **no** son cuasi insesgadas y que por ende los gráficos de control a los que pertenecen deberían ser considerados como no óptimos.

Ahora bien, con el fin de caracterizar el ARL_{BSL} del GC u Shewhart, se llevó a cabo un estudio para una amplia gama de “porcentajes de defectos por UDI⁸ (u)” y “tamaños de muestra (n)”. Para esto se desarrolló en R un algoritmo capaz de computar el ARL_{BSL} para cualquier valor de u y n . El estudio se lo hizo para $u = 1(0.1)5$ con rangos de tamaños de muestra específicos para cada u . Además, también se calcularon los ratios de las colas alfa (R_α). Ejemplos de los resultados se muestran en la Figura 59 y en la Figura 60 y también de las figuras incluidas en el Anexo 9.

Para determinar el rango de tamaños de muestras (o rango de n), se aplicaron los siguientes criterios: i) que con todas las combinaciones de u y n se debían obtener límites de control

⁸ UDI = Unidad de inspección

inferior (LCI) y superior (LCS) y ii) que el tamaño de muestra debía ser lo suficientemente grande. Para cumplir con estos criterios se utilizó la ecuación (23) con $d = 0.97u$ para el valor n mínimo y $d = u/2$ para el n máximo. En los gráficos el rango de tamaños de muestra se denota como: n rango = n mínimo (0.1) n máximo, por ejemplo en la Figura 59 el n rango = $10(0.1)36$, indicando que los cálculos fueron realizados desde el tamaño de muestra 10 hasta el 36 en incrementos de 0.1.

Tomando en cuenta los criterios de la Tabla 11, de los resultados se puede concluir lo siguiente:

1. Que el GC u Shewhart produce gráficos de control con ARL_{BSL} negativos, lo que denota que las curvas ARL de estos gráficos tendrán sesgo negativo, lo que a su vez implica que estos GCs tendrán una reducida capacidad para detectar mejoras en el proceso.
2. Que el ARL_{BSL} en los GCs u Shewhart fluctúa por debajo de cero, y que su valor específico depende de los parámetros u y n .
4. Que las curvas ARL de los GCs u Shewhart se acercan al umbral $ARL_{BSL} > -2$ (de curvas ARL cuasi-inesgadas) solo para tamaños de muestra extremadamente altos.

De lo expuesto en las conclusiones 1, 2 y 3 se deduce que con los gráficos de control u Shewhart, no es posible monitorizar procesos de una forma adecuada.

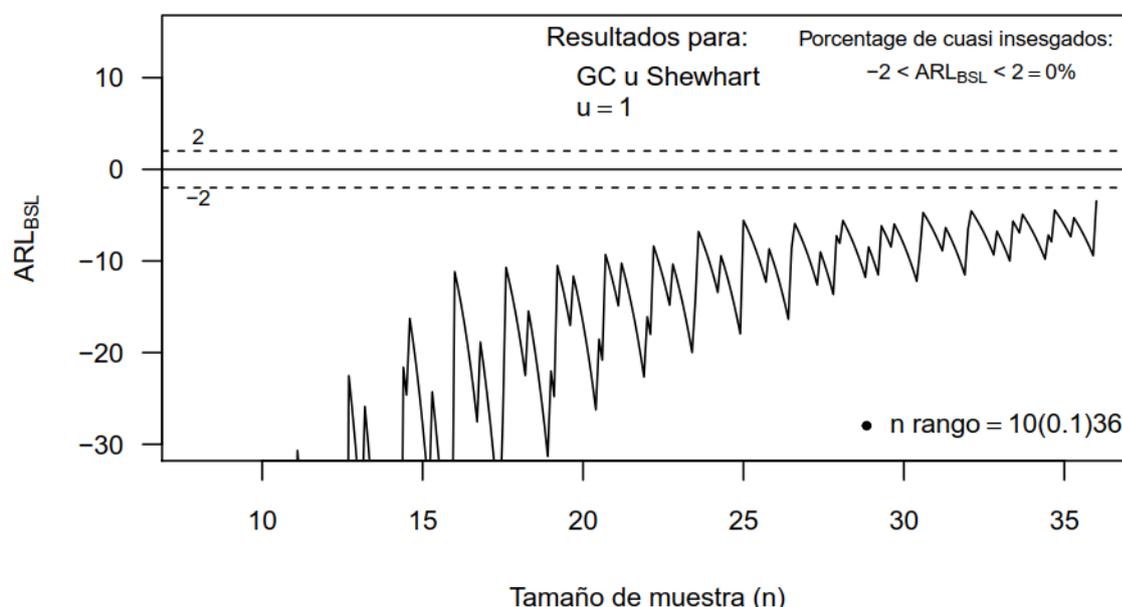


Figura 59. GC u Shewhart – Comportamiento del ARL_{BSL} para $u = 1$

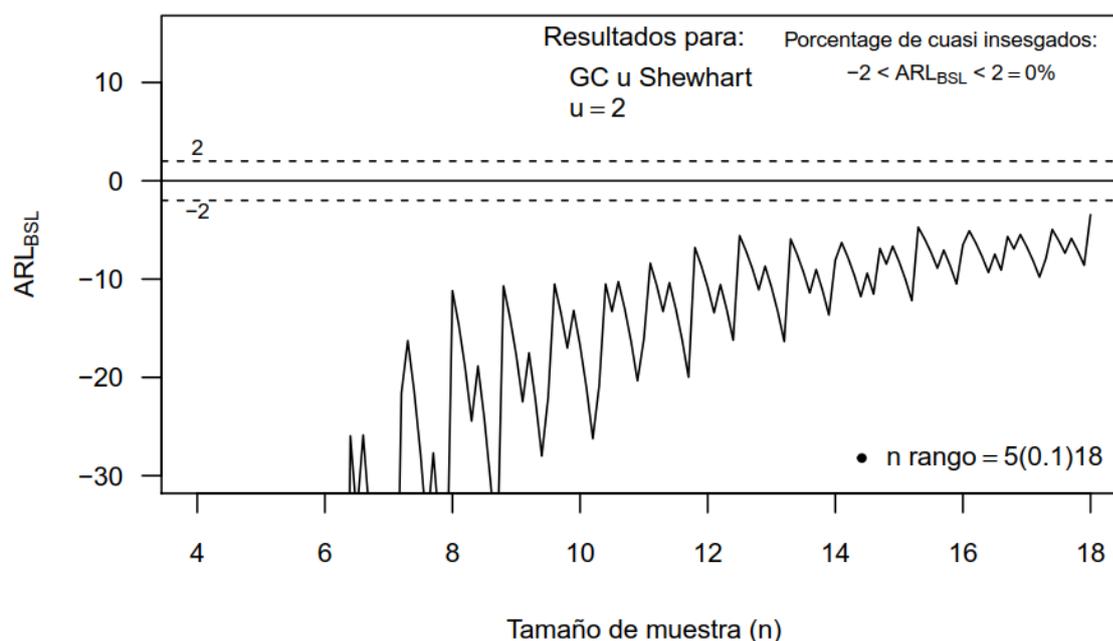


Figura 60. GC u Shewhart – Comportamiento del ARL_{BSL} para $u = 2$

5.6.1.1 Gráfico de control u Shewhart: relación entre las colas alfa y el ARL_{BSL}

Utilizando los resultados obtenidos en el estudio descrito en la sección anterior (ver 5.6.1), se creó la Figura 61 que muestra la relación que existe entre las colas alfa y el ARL_{BSL} . Cabe recalcar que aunque los resultados que se muestran son solo para $u = 1$, el comportamiento es similar desde $u = 1$ hasta $u = 5$ (que es todo el rango de valores evaluado).

Antes de analizar los resultados hay que recordar que el ratio de las colas alfa es igual a $R_\alpha = \alpha_L / \alpha_U$ (ver sección 4.1.7.1) y que este sirve para determinar la disparidad que existe entre las colas alfa. Recordemos también que la situación ideal es que estas sean similares, es decir que $R_\alpha \approx 1$, y que mientras más se aleje el R_α de 1 mayor será la disparidad entre las colas.

Analizando los resultados expuestos en la Figura 61 es posible concluir lo siguiente:

1. En términos generales, cuanto más grande sea la disparidad de las colas alfa, mayor será la severidad del sesgo de las curvas ARL.
2. Que para obtener valores ARL_{BSL} cercanos a cero es necesario reducir, a lo mínimo posible, la disparidad entre las colas alfa.

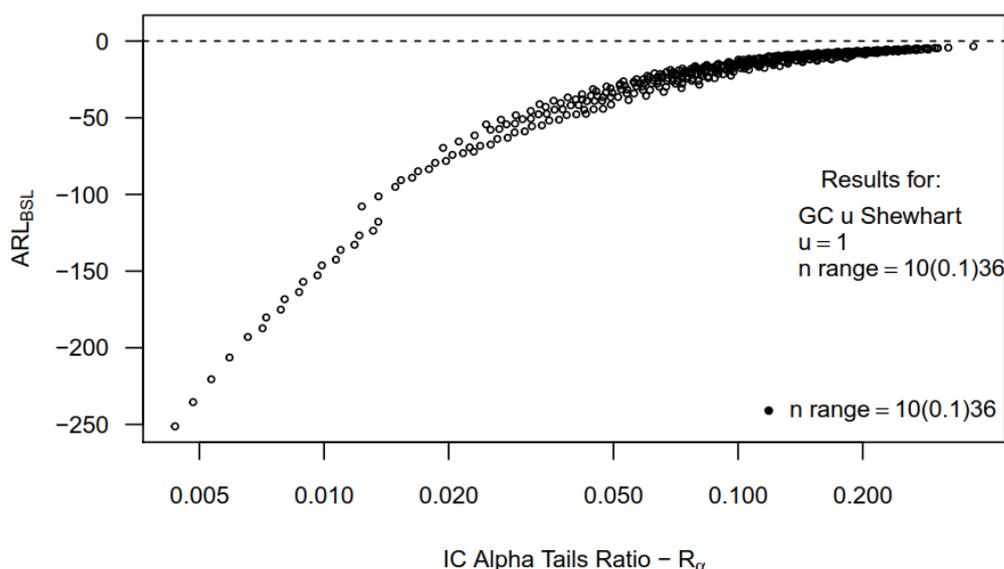


Figura 61. GC u Shewhart - Relación entre R_α y ARL_{BSL} para $u = 1$

5.6.1.2 Sinopsis de la sección 5.6.1

Al final de la sección 3.2.2.2 se mencionó que la pregunta que había dado lugar a la investigación resumida en esta tesis era: *¿Cómo afecta la imprecisión de la aproximación normal a la capacidad de monitorización de los CGs por atributos tipo Shewhart?*

Para obtener una respuesta a esa pregunta, y para el caso específico de GCs por atributos basados en la distribución de Poisson, se desarrolló todo lo incluido en las secciones 4.2 y 5.6.1. Por lo que finalmente, y fundamentado en lo expuesto, es posible afirmar que la imprecisión del ajuste de la aproximación normal a la distribución de Poisson, afecta de forma muy adversa la capacidad de monitorización del GC u Shewhart. Por esta razón la recomendación que se da en esta tesis es que este GC no sea utilizado, especialmente si se requiere detectar mejoras en el proceso.

5.6.2 Gráficos de control u alternativos: Caracterización del ARL_0 y ARL_{BSL}

Una vez establecido que el GC u Shewhart producía gráficos con una capacidad de monitorización no óptima (ver sección 5.6.1), el trabajo se enfocó en cumplir con el tercer objetivo específico de esta tesis, que fue: *Caracterizar la capacidad de monitorización de gráficos de control por atributos relevantes que hayan sido propuestos hasta la fecha de realización de la investigación.*

Para lograr ese objetivo, y para el caso específico de GCs basados en la distribución de Poisson, se efectuó una búsqueda de gráficos de control u que hubiesen sido propuestos como alternativas superiores al GC u Shewhart y que fuesen igual de fáciles de aplicar y utilizar. La

búsqueda trajo a la luz varios GCs entre los cuales se destacan los propuestos por Ryan&Schwertman (1997), Winterbottom (1993) y Kittlitz (2006).

Cabe mencionar que los autores Chen&Cheng (1998) analizaron dos GCs u basados en transformaciones, el primero basado en la transformación Q (Quesenberry, 1991-2) y el segundo basado en la transformación raíz cuadrada (Ryan T. P., 2011, pág. 204), llegando a comprobar que el GC u propuesto por Winterbottom era superior a esos dos GCs. Por este motivo se decidió omitirlos de la lista de GCs u alternativos.

La Tabla 19 contiene los tres GCs u seleccionados. Todos ellos fueron sometidos a un estudio que tuvo como fin caracterizar sus ARL_0 y ARL_{BSL} , y también determinar sus rendimientos en cuanto a la proporción de gráficos de control óptimos que producen dentro de un rango de tamaños de muestra. Los estudios fueron realizados siguiendo los mismos lineamientos de rango de u y de n , que los utilizados para el GC u Shewhart en la sección 5.6.1. En las siguientes secciones se presentan los detalles de cada uno de estos GCs junto con los resultados.

Tabla 19: Gráficos de control u alternativos

<i>Descripción del GC</i>	<i>Autores y publicación</i>
GC u con límites computados a través de una ecuación basada en análisis de regresión.	Ryan y Schwertman (1997)
GC u con límites computados mediante una ecuación basada en la expansión Corner-Fisher	Winterbottom (1993)
GC u cuyos límites se basan en una transformación que el autor llama “Power Transformación”	Kittlitz (2006)

5.6.2.1 Gráfico de control u basado en la expansión Corner-Fisher

Winterbottom (1993) fue quien originalmente propuso el GC u basado en la expansión Corner-Fisher. Años más tarde los autores Chen&Cheng (1998) lo volvieron a presentar añadiendo variaciones a su aplicación. Los límites de control para este gráfico se calculan mediante la ecuación (34).

$$\text{Límites del GC } u \text{ Corner - Fisher} = u \pm K \sqrt{\frac{u}{n} + \frac{4}{3n}} \quad (34)$$

Donde:

- Para el LCS usar: +
- Para el LCI usar: -
- Para límites 3 sigma usar: $K=3$
- Línea central del gráfico = u

La Figura 62 muestra un ejemplo de las colas alfa típicas que se obtienen con los límites computados con (34). Comparando estas colas alfa con las del GC u Shewhart, es fácil darse cuenta que en este caso las colas alfa oscilan alrededor de 0.00135, y no por encima o por debajo, como es el caso para el GC Shewhart.

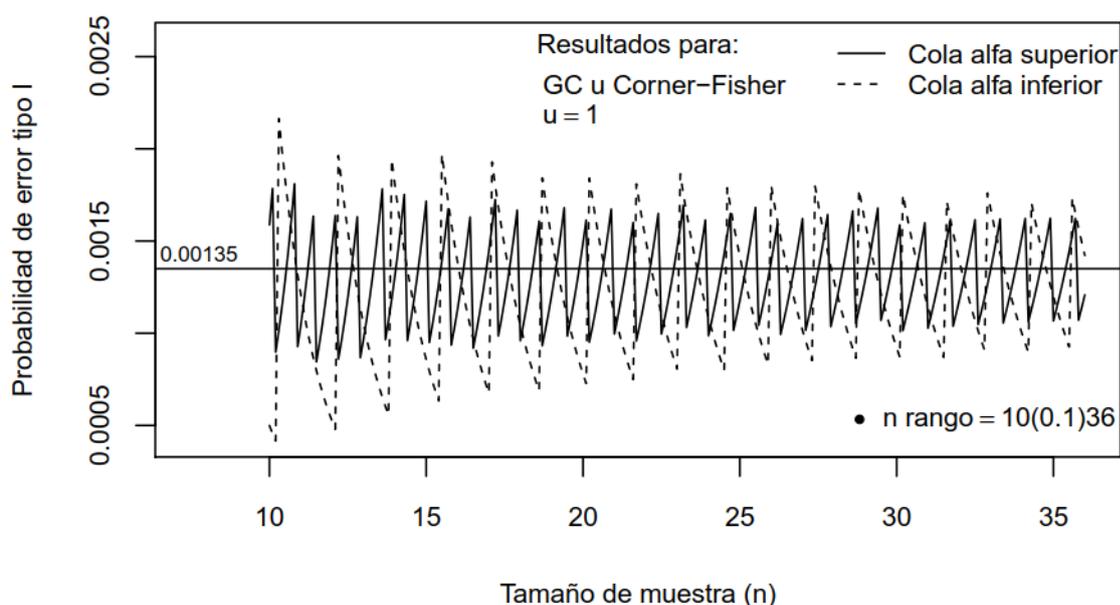


Figura 62. Ejemplo de colas alfa típicas para el GC u Corner-Fisher

La Figura 63 muestra el comportamiento típico del ARL_0 , como se puede apreciar para tamaños de muestra bajos el ARL_0 es, en muchos casos, excesivamente alto y de todas las posibles combinaciones, con $u = 1$ y n rango = 10(1)36, solo el 74% tienen un ARL_0 aceptable (entre 250 y 450).

Los valores ARL_0 excesivamente altos se dan cuando las dos colas están muy por debajo de 0.00135, recordemos que $ARL_0 = 1/\alpha$ donde $\alpha = \alpha_I + \alpha_S$.

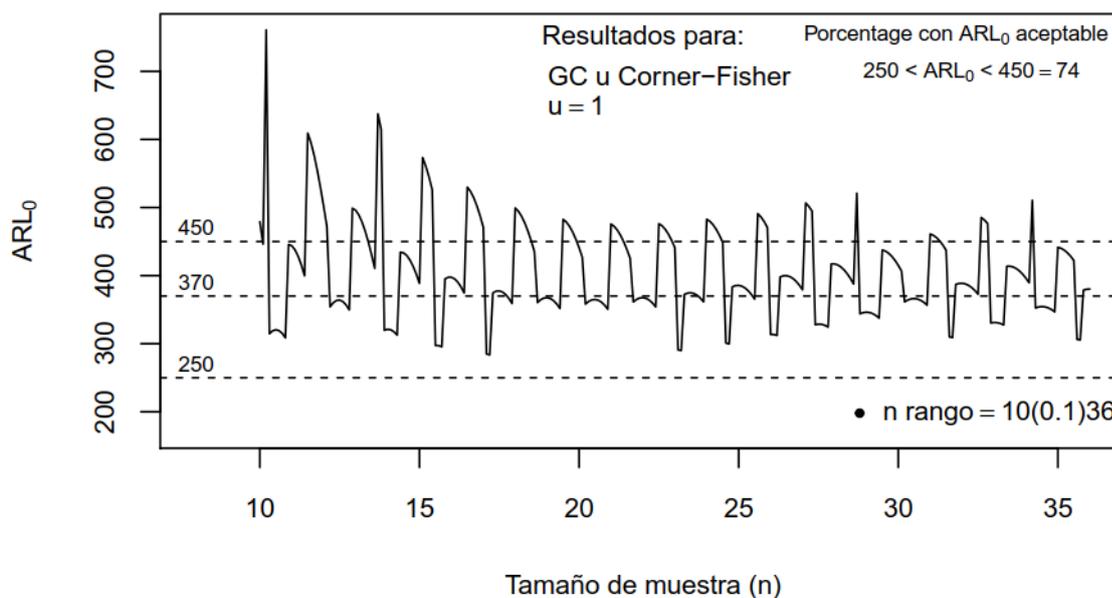


Figura 63. Ejemplo del comportamiento típico del ARL_0 para el GC u Corner-Fisher

La Figura 64 muestra el comportamiento típico del ARL_{BSL} para el GC u Corner-Fisher, comparando estos resultados con los obtenidos para el GC u Shewhart, es fácil deducir que existe una gran mejora. Sin embargo, los resultados también demuestran que del total de combinaciones, con $u = 1$ y n rango = $10(0.1)36$, solo el 77% resultarían en gráficos de control con curvas ARL cuasi insesgadas y que los gráficos con curvas sesgadas se darían principalmente con tamaños de muestra pequeños.

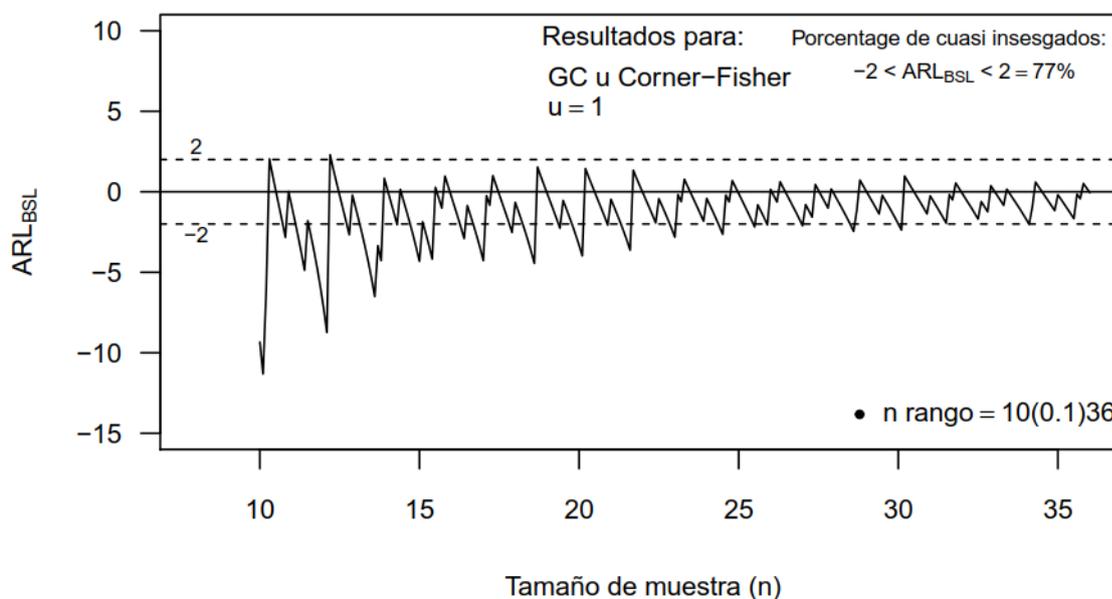


Figura 64. Ejemplo de comportamiento típico del ARL_{BSL} para el GC u Corner-Fisher

La Tabla 20 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC u Corner-Fisher, antes de analizarlos es necesario

enfatar que un gráfico de control puede ser considerado como óptimo solo cuando tiene: i) una curva ARL cuasi incesgada y ii) un ARL_0 aceptable.

Los resultados demuestran que en promedio solo un 76% de gráficos construidos con el GC u Corner-Fisher tendrían curvas ARL cuasi incesgadas y que además, en promedio, de ellos solo el 80% tendrían valores ARL_0 aceptables. De esto se concluye que si se utilizase el GC u Corner-Fisher, habría un alto riesgo de que el gráfico empleado fuese no óptimo.

Tabla 20: Síntesis de los resultados para el GC u Corner-Fisher

u	Rango de tamaños de muestra	Porcentaje de gráficos cuasi incesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi incesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
1	10(0.1)36	77%	80%
2	5(0.1)18	76%	81%
4	2(0.05)9	73%	82%
5	2(0.05)7	76%	77%

Cabe mencionar que Cavalcanti&Cordeiro (2006) presentan un GC u basado en una pequeña variación al propuesto por Winterbottom. Ellos utilizan los cuatro primeros términos de la expansión Corner-Fisher, a diferencia de Winterbottom que utiliza solo los tres primeros. Sin embargo, un estudio comparativo del ARL_{BSL} y ARL_0 de estos dos GCs, demostró que el GC de Winterbottom da mejores resultados que el propuesto por Cavalcanti&Cordeiro.

5.6.2.2 Gráfico de control u basado en análisis de regresión

Ryan y Schwertman (1997) propusieron un GC u basado en una fórmula derivada por medio de análisis de regresión y que fue diseñada para computar límites de control equiparables a los límites 3-sigma del GC u Shewhart.

Los límites de control para el GC u R&S se calculan mediante la ecuación (35).

$$\text{Límites del GC } u \text{ R\&S} = \frac{1}{n} [a + b(nu) + c\sqrt{nu}] \quad (35)$$

Donde:

- Para el LCS usar: $a=0.6195$; $b=1.0052$; $c=2.983$
- Para el LCI usar: $a=2.9529$; $b=1.01956$; $c=3.2729$
- Línea central del gráfico = u

La Figura 65 muestra un ejemplo de las colas alfa que se obtienen con los límites computados con (35), mientras que la Figura 66 muestra el comportamiento típico del ARL_0 , como se puede apreciar para tamaños de muestra pequeños el ARL_0 es, en muchos casos,

excesivamente bajo y de todas las posibles combinaciones, con $u = 1$ y n rango = $10(0.1)36$, solo el 33% tendrían un ARL_0 aceptable (entre 250 y 450).

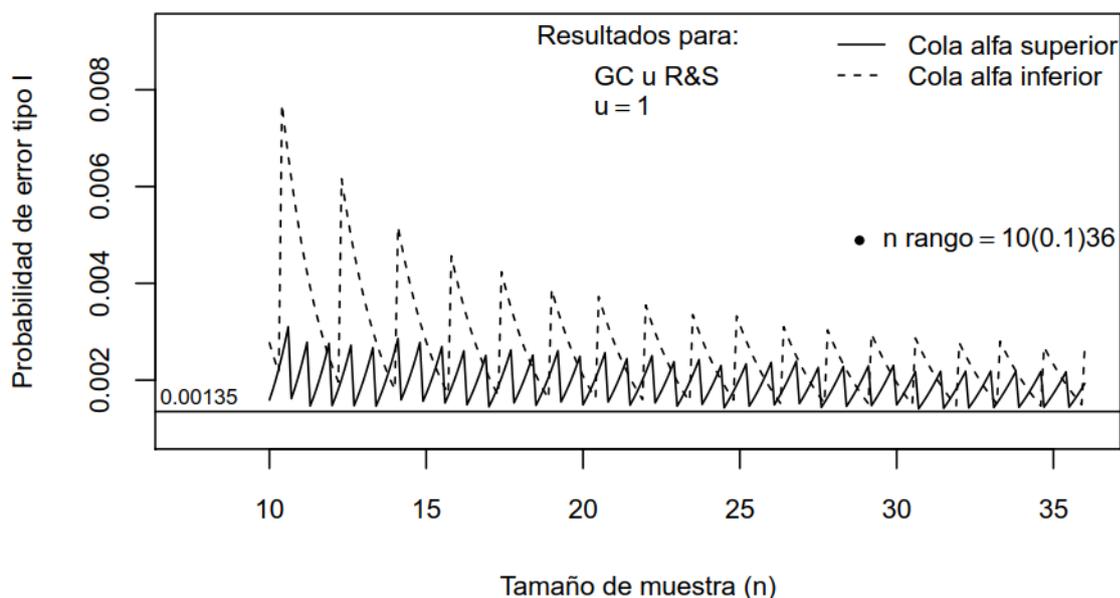


Figura 65. Ejemplo de colas alfa típicas para el GC u R&S

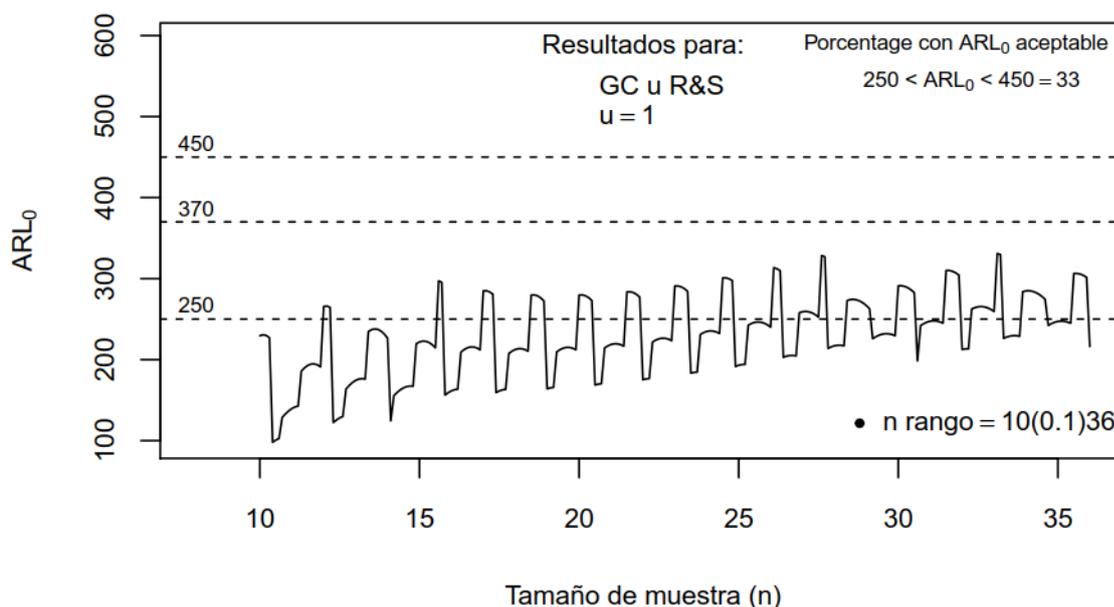


Figura 66. Ejemplo del comportamiento típico del ARL_0 para el GC u R&S

La Figura 67 muestra el comportamiento típico del ARL_{BSL} para el GC u R&S, si se comparan estos resultados con los obtenidos para el GC u Shewhart es fácil deducir que existe una gran mejora. Notese que del total de GCs que se obtienen con las combinaciones $u = 1$ y n rango = $10(0.1)36$, un 90% son gráficos con curvas ARL casi inesgadas.

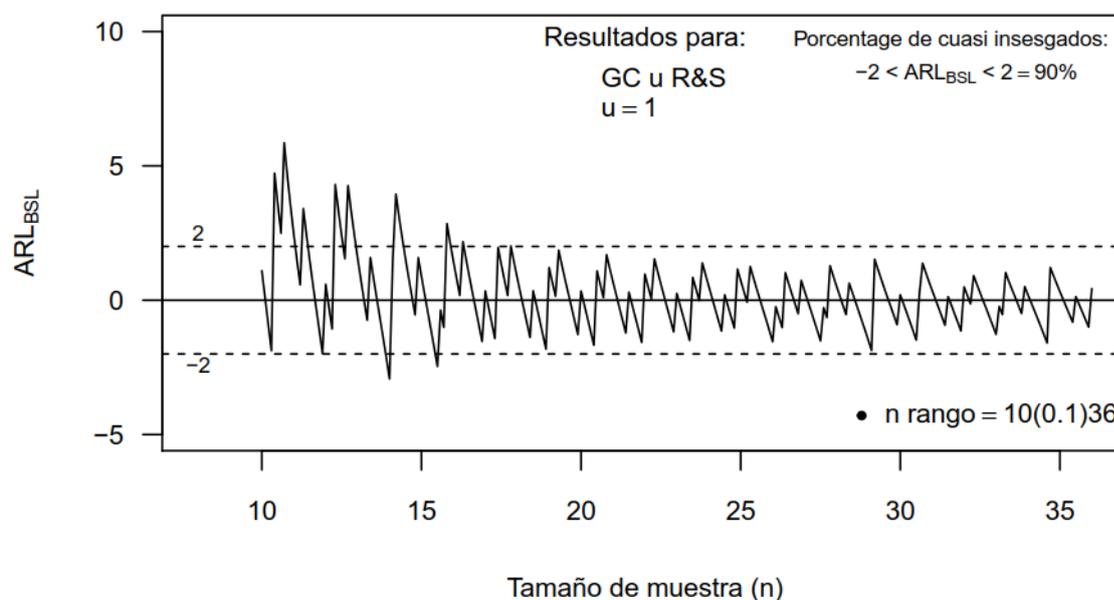


Figura 67. Ejemplo del comportamiento típico del ARL_{BSL} para el GC u R&S

La Tabla 21 contiene una síntesis de los resultados para el GC u R&S. Los resultados demuestran que una alta proporción de los gráficos que se obtienen con este GC tendrían curvas cuasi incesgadas, pero que en cambio solo una baja proporción de ellos tendrían valores ARL_0 aceptables. De esto se concluye que si se utilizase el GC u R&S, habría un alto riesgo de trabajar con un gráfico que fuese no óptimo.

Tabla 21. Síntesis de los resultados para el GC u R&S

u	Rango de tamaños de muestra	Porcentaje de gráficos cuasi incesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi incesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
1	10(0.1)36	90%	37%
2	5(0.1)18	91%	35%
4	2(0.05)9	88%	34%
5	2(0.05)7	89%	37%

5.6.2.3 Gráfico de control u basado en la transformación “Power Transformación”

Kittlitz (2006) propuso un GC basado en una transformación llamada “Power Transformation”. La ecuación que da Kittlitz computa límites de control para gráficos de control c , sin embargo con una sencilla modificación esa ecuación puede ser adaptada para

computar límites de gráficos u . Los límites de control para el GC u PT se calculan mediante la ecuación (36).

$$\text{Límites del GC } u \text{ PT} = \frac{1}{n} \left\{ \left[\left(C + \frac{1}{12} \right)^{2/3} \pm K \left(\frac{2}{3} \right) C^{1/6} \right]^{3/2} + a \right\} \quad (36)$$

Donde:

- $C = un$
- Para el LCS usar: $+ y a = -\frac{3}{4}$
- Para el LCI usar: $- y a = \frac{1}{4}$
- Para límites 3 sigma usar: $K=3$
- Línea central del gráfico = u

La Figura 68 muestra un ejemplo de las colas alfa que se obtienen con los límites computados con (36). Mientras que la Figura 69 muestra el comportamiento típico del ARL_0 , como se puede apreciar el ARL_0 es, en muchos casos, excesivamente bajo y de todas las posibles combinaciones, con $u = 1$ y n rango = 10(0.1)36, solo el 53% tendrían un ARL_0 aceptable (entre 250 y 450).

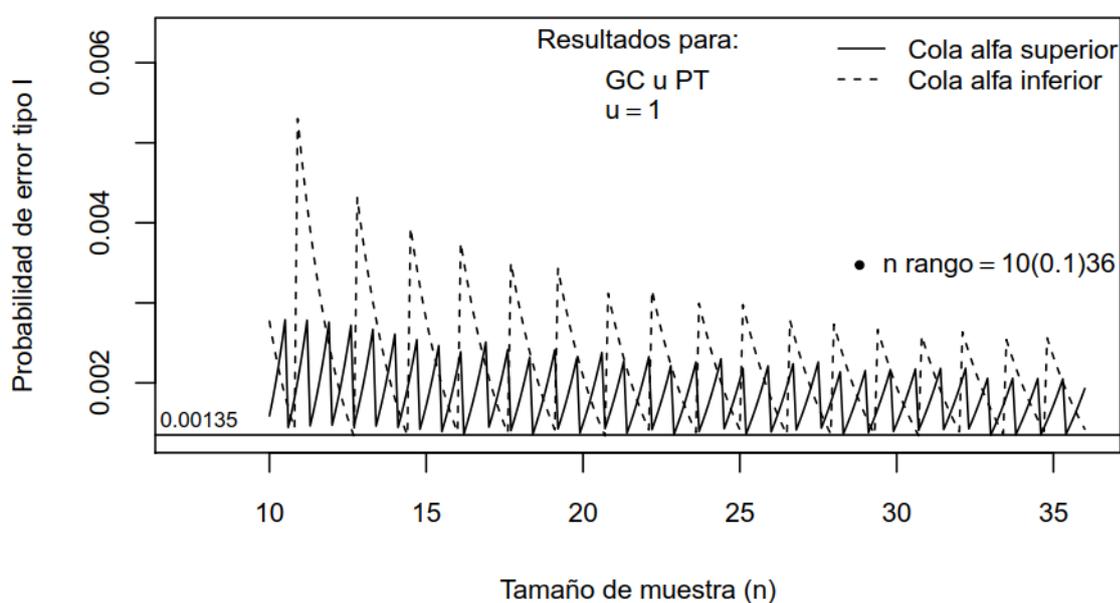


Figura 68. Ejemplo de colas alfa típicas para el GC u PT

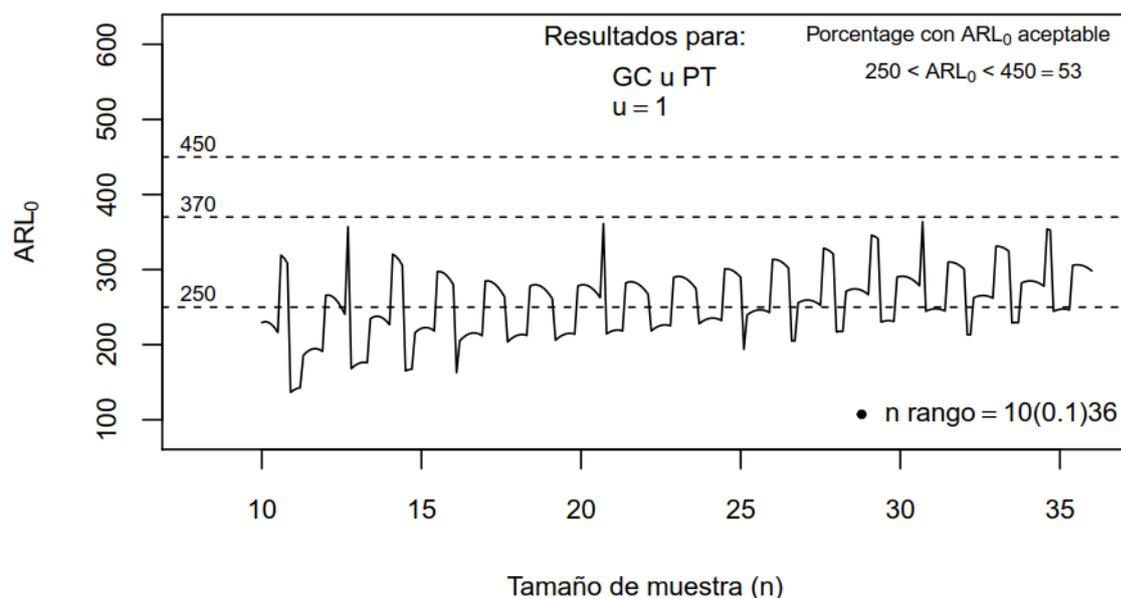


Figura 69. Ejemplo del comportamiento típico del ARL_0 para el GC u PT

La Figura 70 muestra el comportamiento típico del ARL_{BSL} para el GC u PT, si se comparan estos resultados con los obtenidos para el GC u Shewhart, es evidente que existe una gran mejora. Notese que del total de combinaciones, con $u = 1$ y n rango = 10(0.1)36, el 87% serían gráficos de control con curvas ARL cuasi insesgadas.

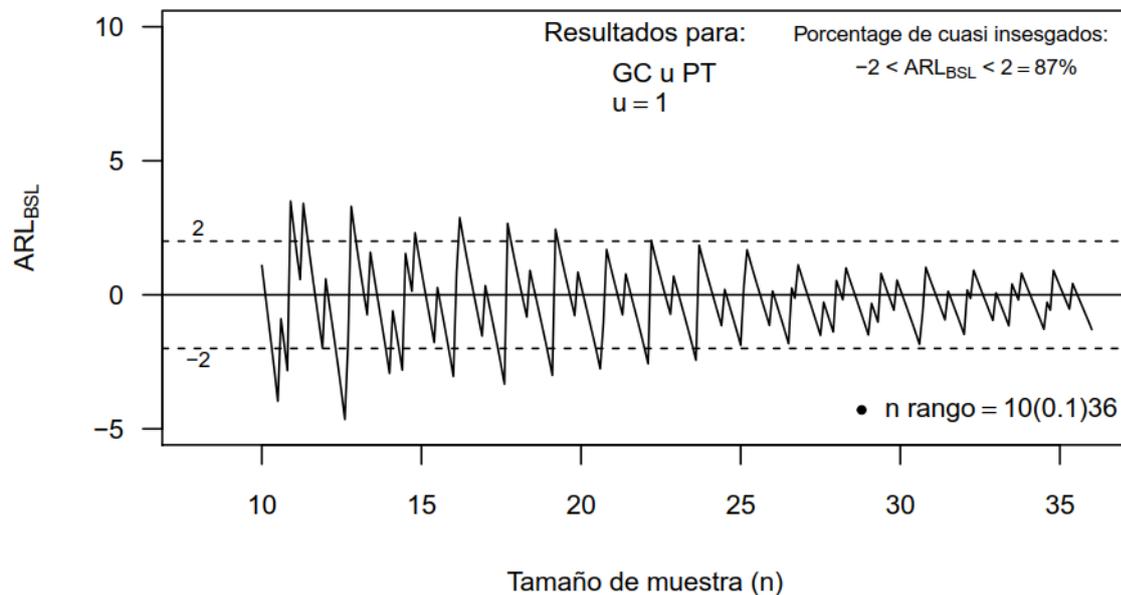


Figura 70. Ejemplo del comportamiento típico del ARL_{BSL} para el GC u PT

La Tabla 22 contiene una síntesis de los resultados para el GC u PT. Los resultados demuestran que una proporción significativa de gráficos construidos con este GC tendrían curvas cuasi insesgadas. Sin embargo, los resultados del ARL_0 son muy pobres. De esto se

concluye que si se utilizase el GC u PT, habría un alto riesgo de trabajar con un gráfico no óptimo.

Tabla 22. Síntesis del rendimiento ARL para el GC u PT

u	Rango de tamaños de muestra	Porcentaje de gráficos cuasi incesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi incesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
1	10(0.1)36	87%	54%
2	5(0.1)18	84%	54%
4	2(0.05)9	84%	52%
5	2(0.05)7	84%	54%

5.6.2.4 Sinopsis de la sección 5.6.2

Los resultados obtenidos en la fase de investigación resumida en la sección 5.6.2, demuestran que existen varios GC u alternativos que son superiores al GC u Shewhart. Sin embargo, una conclusión muy importante derivada de los resultados expuestos en esta sección, es que estos GCs alternativos también pueden tener curvas ARL sesgadas y/o ARL_0 no aceptables. Por lo que se deduce que si se utilizara cualquiera de estos GCs, se correría el riesgo de trabajar con gráficos del control no óptimos.

5.6.3 El gráfico de control u Kmod

En esta sección se resume el trabajo realizado para cumplir con el cuarto objetivo específico de la investigación que fue “Desarrollar nuevas propuestas de gráficos de control por atributos”. Para lograr este objetivo se creó el GC u Kmod, que sirve para monitorizar procesos mediante la distribución de Poisson. Este GC tiene un rendimiento, en cuanto a la proporción de gráficos idóneos, superior a todos los GCs u alternativos analizados anteriormente y además cuenta con un método muy sencillo para verificar si sus curvas ARL son o no “cuasi incesgadas”.

5.6.3.1 El factor K adaptativo y los límites de control para el gráfico de control u Kmod

En la sección 5.6.1.1 se concluyó que para obtener valores ARL_{BSL} cercanos a cero es necesario reducir, a lo mínimo posible, la disparidad entre las colas alfa. Tomando en cuenta este antecedente, y en base a lo realizado para el GC p Kmod (ver sección 5.5.3.1), se encontró que muy buenos resultados se obtenían añadiendo el término $(T_{1|S})/\sqrt{un}$ al factor K

= 3 de la ecuación para el GC u Shewhart mostrada en la Tabla 7 de la sección 4.2.4. Nótese que el factor que se añade a $K = 3$ está basado en la desviación estándar de la distribución de Poisson.

El factor K modificado resultante es $[3 \pm (T_{I|S})/\sqrt{un}]$, donde los signos más y menos están relacionados con los límites de control inferior y superior respectivamente, y el término es una constante que debe ser determinada para cada límite de control (T_I para el límite inferior y T_S para el superior). A este factor K modificado se lo ha denominado K_{mod} , ver ecuación (37).

Para determinar valores óptimos de T_I y T_S se realizó un estudio de optimización heurística basado en los siguientes criterios:

- 1 Ratio de las colas alfa lo más cercanos a 1
- 2 ARL_{BSL} lo más cercano a cero
- 3 ARL_0 entre 250 y 450

El estudio consistió en primeramente dar valores entre $0.8(0.1)2$ a cada una de las constantes T_I y T_S de la ecuación (37). Luego de eso, y para cada uno de esos valores, se computaron individualmente las colas alfa (inferior y superior) que se obtenían para $u = 1(0.5)5$ usando n rangos específicos para cada u , y cuyos valores mínimo y máximo fueron determinados siguiendo los mismos lineamientos utilizados para el análisis de las colas alfa del GC u -Shewhart (ver sección 4.2.5.1). Una vez obtenidos los resultados de las colas alfa, se identificaron los valores más destacados de T_I y T_S , es decir aquellos con los que, en promedio y para cada u , se obtenían colas alfa lo más cercanas a 0.00135. Una vez identificados los valores T_I y T_S destacados, se procedió a combinarlos metódicamente en (37) para luego, y con cada una de estas combinaciones, computar los R_α , ARL_{BSL} y ARL_0 para los mismos valores de u y n mencionados anteriormente. Una vez hecho esto, y con base en los criterios de optimización descritos previamente, se analizaron los resultados hasta dar con los valores óptimos de T_I y T_S .

Luego de un extenso análisis de los resultados mencionados al final del párrafo anterior, se determinó que con los valores $T_S = 1.2$ para el K_{mod} del LCS y $T_I = 1.7$ para el K_{mod} del LCI se obtenían los mejores resultados. La ecuación (37) sirve para computar los límites de control del nuevo gráfico de control al que se lo ha denominado como u K_{mod} .

$$GC\ u\ K_{mod}\ \text{Límites de control} = u \pm (K_{mod}_{LCI|LCS}) \sqrt{\frac{u}{n}} \quad (37)$$

Dónde:

Para el limite de control superior (LCS):

$$Kmod_{LCS} = 3 + \frac{T_S}{\sqrt{un}} ; T_S = 1.2$$

Para el limite de control inferior (LCI):

$$Kmod_{LCI} = 3 - \frac{T_I}{\sqrt{un}} ; T_I = 1.7$$

Línea central del GC u Kmod = u

5.6.3.2 Gráfico de control u Kmod: Caracterización del ARL_{BSL} y ARL_0

El GC u Kmod fue sometido a estudio que tuvo como fin caracterizar el comportamiento del ARL_0 y ARL_{BSL} , y también determinar su rendimiento en cuanto a la proporción de gráficos de control óptimos que se obtienen dentro de una rango de tamaños de muestra. Para realizar este estudio se utilizaron los mismos lineamientos, para el rango de u y n , que los empleados en la sección 5.6.1 para el GC u Shewhart.

La Figura 71 muestra un ejemplo de las colas alfa típicas que se obtienen con los límites computados con la ecuación (37). Comparando estas colas alfa con las del GC u Shewhart de la Figura 29, es fácil darse cuenta que en promedio las colas alfa del GC u Kmod están más cerca de 0.00135.

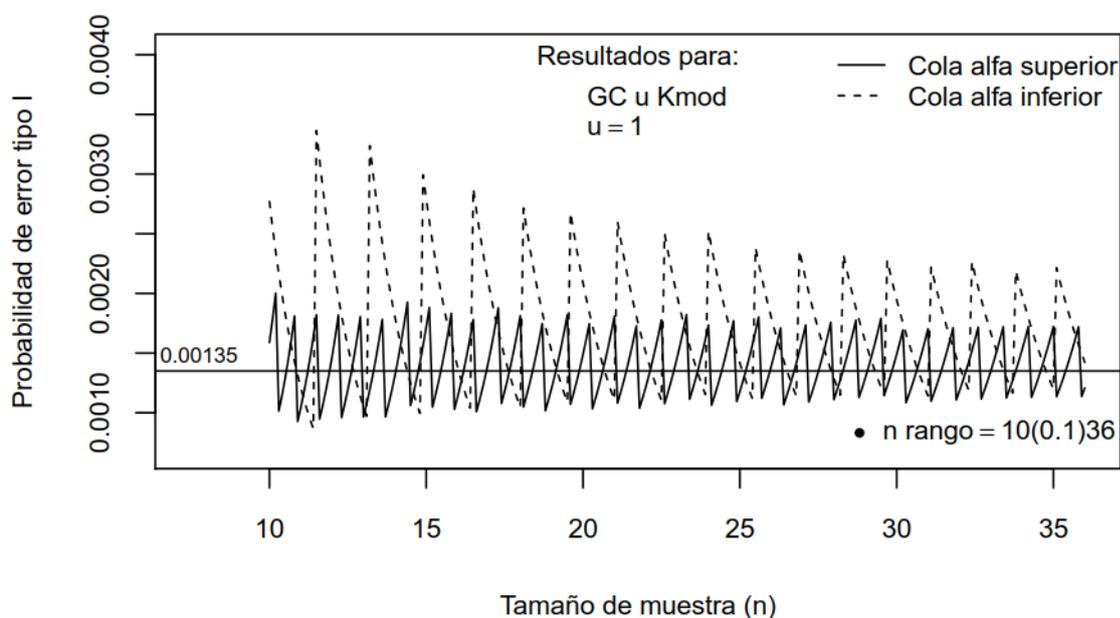


Figura 71. Ejemplo de colas alfa típicas para el GC u Kmod

La Figura 72 muestra el comportamiento del ARL_0 , como se puede observar 93% de los gráficos de control u Kmod contruidos con $u = 1$ y n rango = 10(0.1)36 tendrían valores ARL_0

aceptables. Por otra parte la Figura 73 muestra que 87% de los gráficos en ese rango de n tendrían curvas ARL casi insesgadas.

Resultados del ARL_0 y ARL_{BSL} para $u = 2$ se incluyen en el Anexo 10.

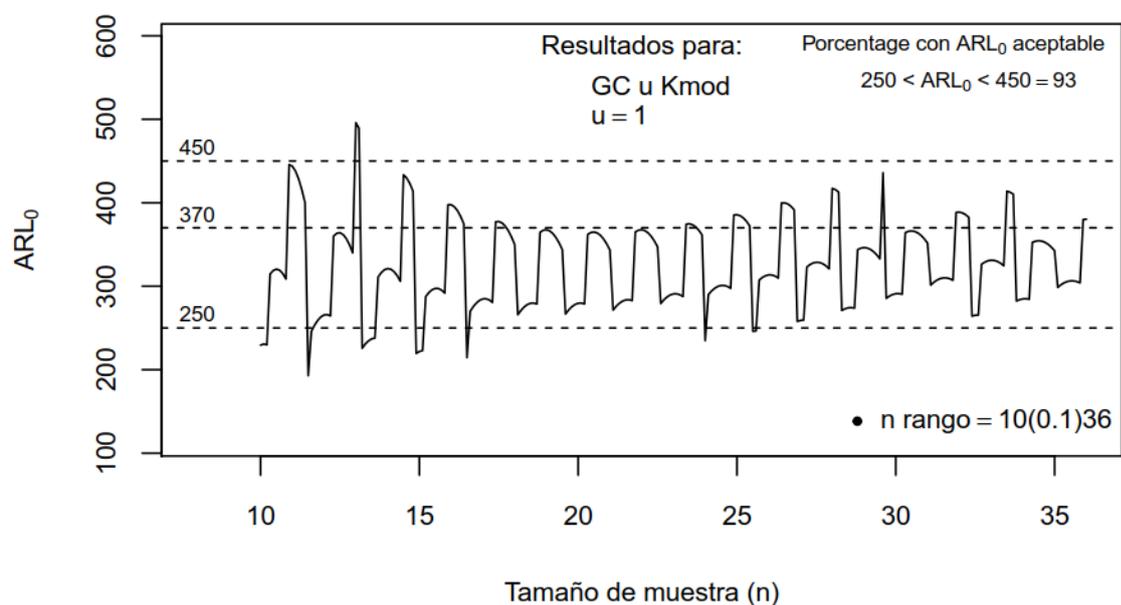


Figura 72. Ejemplo del comportamiento típico del ARL_0 para el GC u Kmod

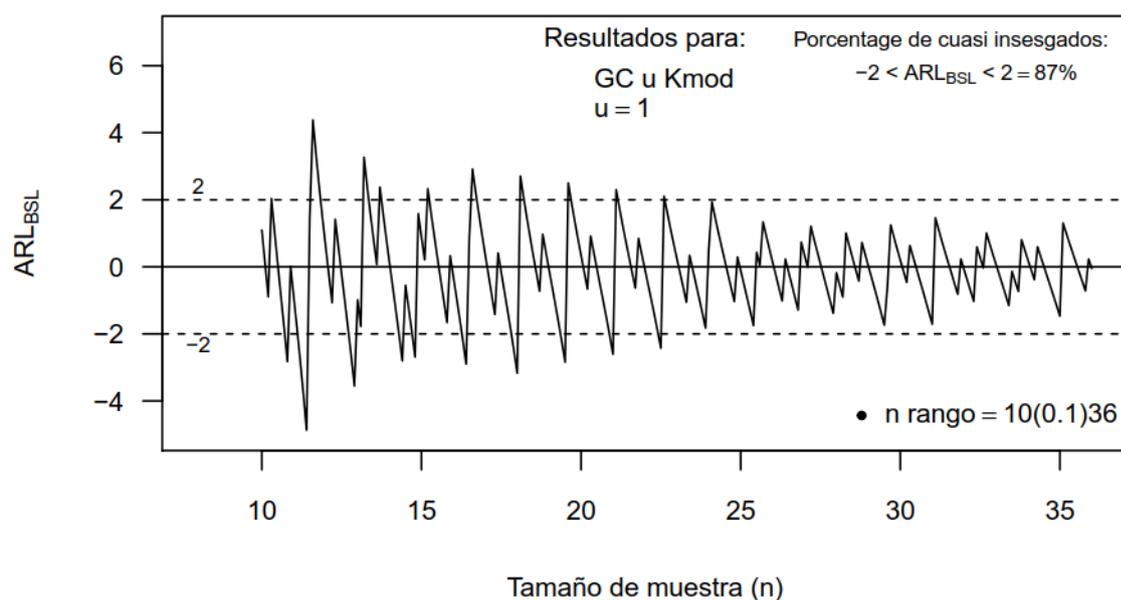


Figura 73. Ejemplo de comportamiento típico del ARL_{BSL} para el GC u Kmod

La Tabla 23 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos (gráficos con curvas ARL casi insesgadas y ARL_0 aceptables) que se obtienen con el GC u Kmod.

Los resultados demuestran que en promedio un 87% de gráficos construidos con el GC u Kmod tendrían curvas ARL cuasi insesgadas y que además, en promedio, un 92% de ellos tendrían valores ARL_0 óptimos. Estas proporciones en términos generales son más altas que las de cualquier otro GC u alternativo incluido en esta tesis.

Tabla 23: Síntesis de los resultados para el GC u Kmod

u	Rango de tamaños de muestra	Porcentaje de gráficos cuasi insesgados: $-2 < ARL_{BSL} < 2$	Porcentaje de gráficos cuasi insesgados con ARL_0 aceptable: $250 < ARL_0 < 450$
1	10(0.1)36	87%	93%
2	5(0.1)18	86%	93%
4	2(0.05)9	84%	92%
5	2(0.05)7	89%	91%

5.6.3.3 Gráfico de control u Kmod: Comparación con otros gráficos de control

La Tabla 24 contiene una síntesis de los resultados obtenidos para la proporción de gráficos óptimos que se alcanzan con el GC u Kmod y con los GCs alternativos más destacados. Antes de comentar acerca de los resultados, cabe recordar que en esta tesis un gráfico de control es considerado como óptimo solo cuando tiene: i) una curva ARL cuasi insesgada y ii) un ARL_0 aceptable.

Tabla 24. Síntesis de los resultados para el GC u Kmod y para los GCs u alternativos más destacados

$GC u$	u	n rango	% de gráficos con curvas ARL dentro de $ARL_{BSL} \pm 2$	Rendimiento del ARL_0 para gráficos que tienen curvas ARL dentro de $ARL_{BSL} \pm 2$					% dentro de $250 < ARL_0 < 450$
				ARL_0 cuartiles (%)					
				0	25	50	75	100	
Kmod	1	10(0.1)36	87%	193	285	320	365	496	93%
RB			90%	125	213	235	274	331	37%
CF			77%	284	357	378	435	609	80%
PT			87%	142	226	261	290	364	53%
Kmod	2	5(0.1)18	86%	222	287	321	365	496	92%
RB			91%	130	213	232	273	330	35%
CF			76%	284	355	379	432	542	81%

PT			84%	143	230	261	289	354	54%
Kmod	4	2(0.05)9	84%	217	285	319	365	496	92%
RB			88%	119	213	231	272	330	34%
CF			73%	284	354	379	435	524	82%
PT			82%	143	225	259	288	354	52%
Kmod	5	2(0.05)7	89%	193	282	314	362	496	91%
RB			89%	143	213	233	274	329	37%
CF			76%	294	360	380	442	609	77%
PT			84%	165	229	259	285	251	54%

Comparando los resultados de cada GC en la Tabla 24, se puede concluir que el GC u Kmod tiene las proporciones más altas de gráficos cuasi insesgados con ARL_0 aceptables. Por ejemplo para $u = 1$ el porcentaje de cuasi insesgados es 87% y de estos el 93% tendrán un ARL_0 aceptable, ningún otro GC sobrepasa esos porcentajes. Por lo que, basado en los resultados expuestos en esa tabla, es posible concluir que el GC u Kmod es superior a cualquiera de los GCs u alternativos incluidos en esta tesis.

Sin embargo, de los resultados también es posible deducir que con el GC u Kmod todavía existe el riesgo, aunque mucho menor que con los otros GCs, de obtener gráficos de control con curvas ARL sesgadas y/o ARL_0 no aceptables. Como se ha dicho anteriormente, computar el ARL_0 para verificar si este es aceptable es relativamente sencillo ya que solo se necesitan computar los valores de las colas alfa, sin embargo computar el ARL_{BSL} es mucho más complejo. Con esto en mente, y basados en el método presentado en la sección 5.5.3.4, se concibió un método que permite determinar, de una manera muy sencilla, si un GC u Kmod tiene o no una curva ARL insesgada. Este método se detalla en la siguiente sección.

5.6.3.4 Gráfico de control u Kmod: Método para determinar si una curva ARL es cuasi insesgada

En esta sección se expone un método que permite verificar si un gráfico GC u Kmod tiene o no una curva ARL cuasi insesgada. Para ilustrar como funciona este método, se incluye la Figura 55 que muestra la relación que existe entre el ARL_{BSL} y el ratio de las colas alfa obtenidas con el GC u Kmod para $u = 1$. Cabe mencionar que con este método se cumple con el segundo objetivo específico de esta tesis.

En esa figura se han trazado los límites $R_\alpha = 0.65$ y $R_\alpha = 2.4$ dentro de los cuales, y como se puede observar, la gran mayoría de combinaciones u y n están dentro del criterio $-2 < ARL_{BSL}$

< 2 para curvas cuasi ARL insesgadas. Estos límites fueron obtenidos luego de un extenso análisis de resultados de ARL_{BSL} vs R_α entre $u = 1$ y $u = 5$. El método para determinar si un GC u Kmod tiene una curva ARL cuasi insesgada o no, consiste simplemente en verificar si el ratio de las colas alfa cae dentro de los límites $0.65 < R_\alpha < 2.4$.

Sin embargo, cabe mencionar que en ocasiones incluso con R_α dentro de 0.65 y 2.4, el ARL_{BSL} para ciertas combinaciones de u y n podría exceder los límites $ARL_{BSL} = \pm 2$. En los análisis realizados se identificó que por lo general cuando esto sucede, los límites excedidos no sobrepasan los valores $ARL_{BSL} = \pm 3$, por lo que el sesgo de las curvas ARL para estos casos, podría considerarse que está dentro en un límite tope de aceptabilidad.

Una recomendación que se da es que cuando se obtengan valores de R_α muy cercanos a 0.65 se determine el ARL_{BSL} , esto debido a que en ese límite pueden obtenerse valores ARL_{BSL} muy por debajo -2.

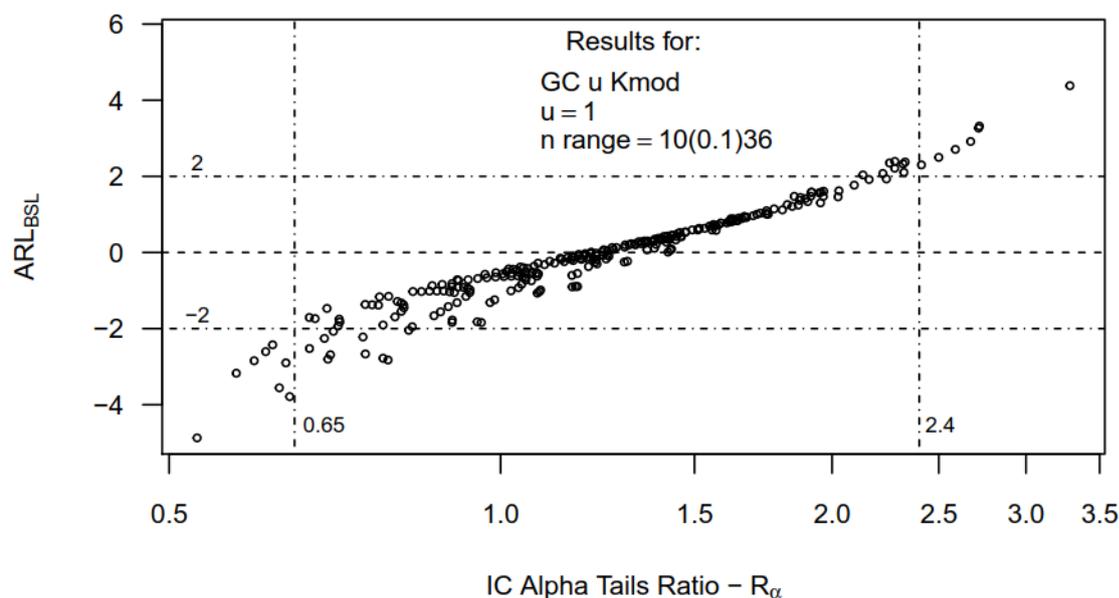


Figura 74. GC u Kmod: Relación entre el ARL_{BSL} y R_α para $u = 1$

5.6.3.5 Gráfico de control u Kmod: ejemplo de curvas ARL

La Figura 75 muestra las curvas ARL obtenidas con el GC u Kmod para $u = 1$ con $n = 11$ y $n = 18$. Estos valores de u y n son los mismos que los utilizados para obtener las curvas ARL de la Figura 31 con el GC u Shewhart (ver sección 4.2.6). Comparando las curvas de esas dos figuras (75 y 31) se puede apreciar que con el GC u Kmod la severidad del sesgo se ha reducido considerablemente.

En las curvas de la Figura 75 se puede observar que la perteneciente a $n = 11$ tiene un $ARL_{BSL} = -0.9$ y un $ARL_0 = 444$, los dos son valores bastante aceptables para una curva ARL.

Mientras que la perteneciente a $n = 18$ tiene un $ARL_{BSL} = -3.13$ lo que indica que esta curva tiene un sesgo inaceptable.

El caso de la curva para $n = 18$ y $u = 1$ sirve de ejemplo para demostrar lo siguiente:

I. El uso del ratio de las colas alfa para identificar si la curva es o no cuasi insegada

Para esta curva el $R_\alpha = 0.58$, lo que está por debajo del límite $R_\alpha = 0.65$ del método descrito en la sección 5.6.3.4. Es decir, con tan solo computar el R_α se podría haber identificado correctamente que el sesgo de la curva ARL para $n = 18$ y $u = 1$ sería inaceptable.

II. Cambiar el tamaño de muestra para obtener ARL_{BSL} y ARL_0 aceptables:

Una recomendación que se da en esta tesis en caso de que se obtengan GCs u Kmod con valores ARL_{BSL} y/o ARL_0 inaceptables, es que se modifique el tamaño de muestra hasta que estos se vuelvan aceptables.

Por ejemplo, para dar solución al caso de $n = 18$ y $u = 1$ se modificó el tamaño de muestra hasta dar con $n = 17.5$ (asumiendo que es posible utilizar 17.5 UDI), con el que se obtienen $R_\alpha = 1.25$ y $ARL_0 = 377$. Es decir, con tan solo reducir el tamaño de muestra de 18 a 17.5, se obtendría un gráfico de control u Kmod idóneo.

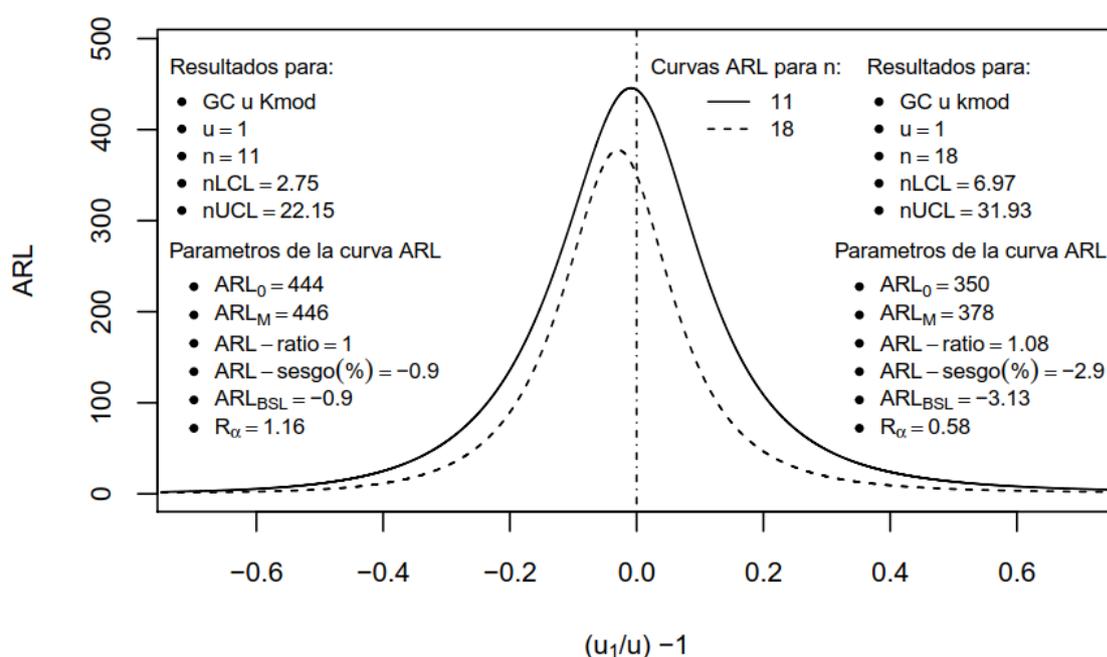


Figura 75. GC u Kmod - Ejemplo de curvas ARL

5.6.3.6 Gráfico de control u Kmod: Aplicación al caso de monitorización de procesos mediante tamaños de muestra variables

En muchas ocasiones es preferible, o más conveniente, monitorizar un proceso mediante tamaños de muestras variables ($n_i : i= 1, 2, 3\dots$). Para este caso, y asumiendo que u es constante, el procedimiento estándar es establecer límites de control para cada n_i , lo que resulta en un gráfico con límites de control variables, ver Montgomery (2009, pág. 319).

Ahora bien, la distribución de Poisson adquiere valores discretos específicos para cada combinación de u y n . Por lo que, si se utilizara tamaños de muestra variables, el valor de las colas alfa variaría y sería específico para cada n_i . Lo que en consecuencia resultaría en valores de ARL_0 y ARL_{BSL} también específicos para cada tamaño de muestra usado.

Nuestra recomendación en este caso es emplear el Kmod u Chart de tal forma que con los n_i usados se obtengan siempre curvas ARL casi insesgadas y ARL_0 aceptables. Esto equivaldría a monitorizar el proceso con una curva ARL casi insesgada cuya capacidad de detección variaría acorde a n_i .

5.7 Aplicación de un gráfico de control Kmod a un caso real

Para ejemplificar el uso de los gráficos Kmod propuestos en esta tesis, se utilizan los datos de la publicación hecha por Mengesha et al (2013), en la cual los autores resumen un trabajo práctico enfocado en mejorar la calidad de una compañía productora de botellas de vidrio.

En esa publicación los autores comentan acerca del uso de varias herramientas de la calidad entre las que se encuentra el gráfico de control p Shewhart. Para construir el gráfico ellos siguieron los lineamientos dados en la norma ISO 8258:1991–E y tomaron 25 muestras conformadas por 192 observaciones cada una ($n = 192$).

Los autores realizaron dos fases de muestreo, con los resultados de la primera fase, y por medio del gráfico de control p Shewhart, demostraron que el proceso bajo estudio estaba fuera de control, por lo que sugirieron a la empresa implementar varias acciones de mejora. Una vez acabadas esas acciones, realizaron la segunda fase de muestreo con cuyos resultados, y nuevamente mediante el gráfico Shewhart, demostraron que el proceso estaba ya bajo control.

En esta tesis se utilizan los resultados reportados para la segunda fase de muestreo, ver Tabla 25, y también se incluye el gráfico p Shewhart que se obtiene con estos datos, ver Figura 76.

Tabla 25. Resultados del muestreo obtenidos en la segunda fase de muestreo

Tamaño de muestra usado (n) = 192									
Los resultados se denotan como: i(número de no conformes) ; i = 1 , 2 , 3 ...									
Ejemplo: 20(22) , indica que en la vienteava muestra hubieron 22 no conformes									
1(39)	2(27)	3(28)	4(19)	5(36)	6(32)	7(32)	8(37)	9(38)	10(39)
11(40)	12(33)	13(35)	14(22)	15(34)	16(36)	17(40)	18(3)	19(39)	20(32)
21(31)	22(18)	23(36)	24(37)	25(40)					

Datos tomados de Mengesha et al (2013)

Analizando los puntos trazados en la Figura 76, y dado que todos caen dentro de los límites de control, se podría asumir que el proceso está bajo control, que es a la conclusión a la que llegaron los autores. Nótese que el gráfico Shewhart tiene un $ARL_{BSL} = -3.95$, lo que indica que tiene una curva ARL con un sesgo negativo un poco severo. La curva ARL se muestra en la Figura 78.

Por otro lado, la Figura 77 muestra el gráfico de control p Kmod obtenido con los datos de la Tabla 25. Como se puede apreciar, a comparación de lo que pasa con el Shewhart, aquí dos puntos (muestras 4 y 22) caen fuera del límite de control inferior, por lo que en este caso no se puede concluir que el proceso está bajo control, por esta razón, y si se hubiese utilizado el gráfico Kmod, muy probablemente se habrían requerido de acciones adicionales para detectar el motivo de la excesiva variación. Nótese que el gráfico Kmod tiene un $ARL_{BSL} = 1.68$, lo que indica que tiene una curva cuasi ARL insesgada. La curva ARL se muestra en la Figura 78.

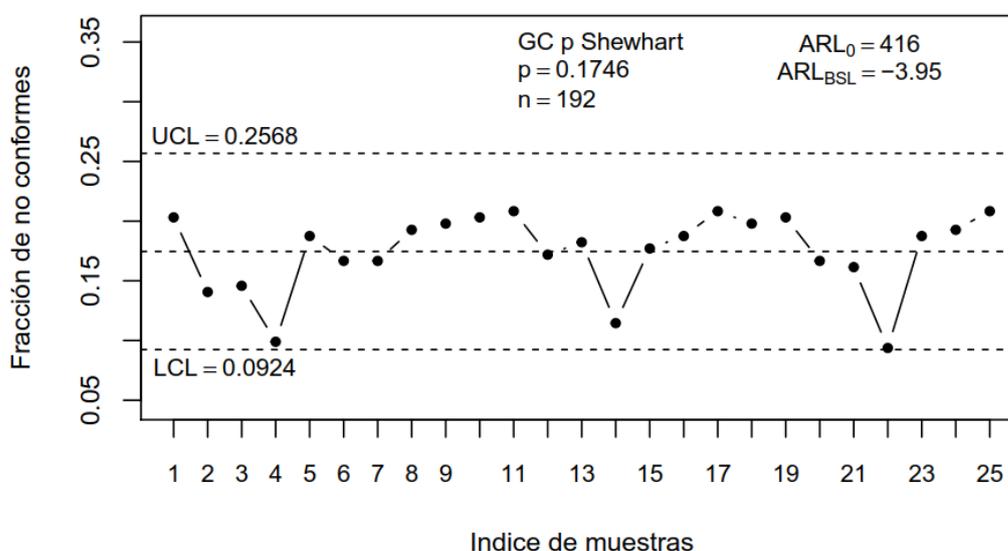


Figura 76. Gráfico de control p Shewhart

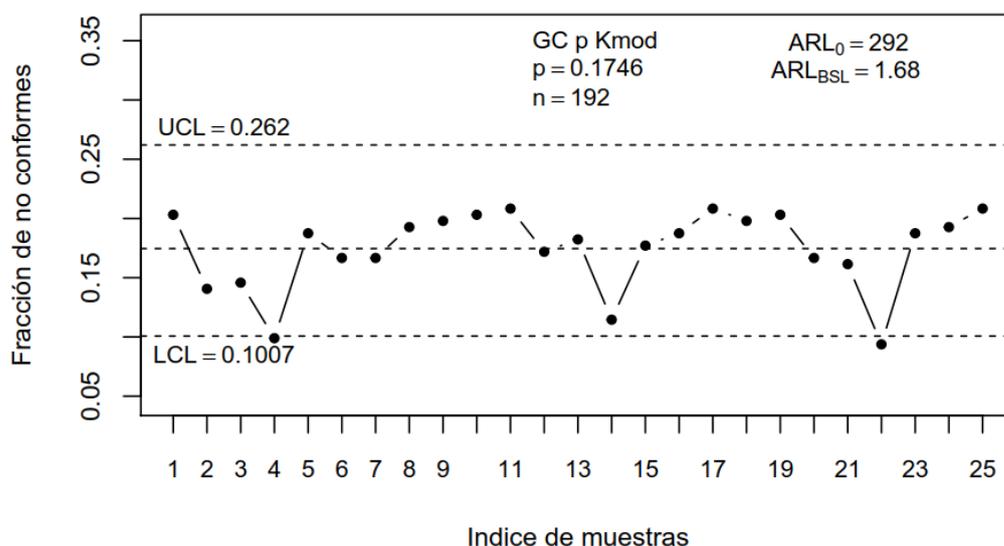


Figura 77. Gráfico de control p Kmod

La Figura 78 muestra las curvas ARL para los gráficos Shewhart y Kmod discutidos en esta sección. Como se puede apreciar la curva del Kmod, al ser cuasi insesgada y tener un valor ARL_0 aceptable, es mucho más idónea que la del Shewhart. Por esta razón es indudable que para este ejemplo el gráfico p Kmod hubiese sido la mejor opción.

Cabe mencionar que si los autores hubiesen conocido acerca del Kmod, hubieran podido determinar si este era idóneo o no con tan solo computar las colas alfa, con las que hubiesen podido determinar que el $ARL_0 = 292$ y, mediante el método descrito en la sección 5.5.3.4, establecer que la curva ARL era óptima debido a que el ratio de las colas está en el límite de $R_\alpha < 2.2$.

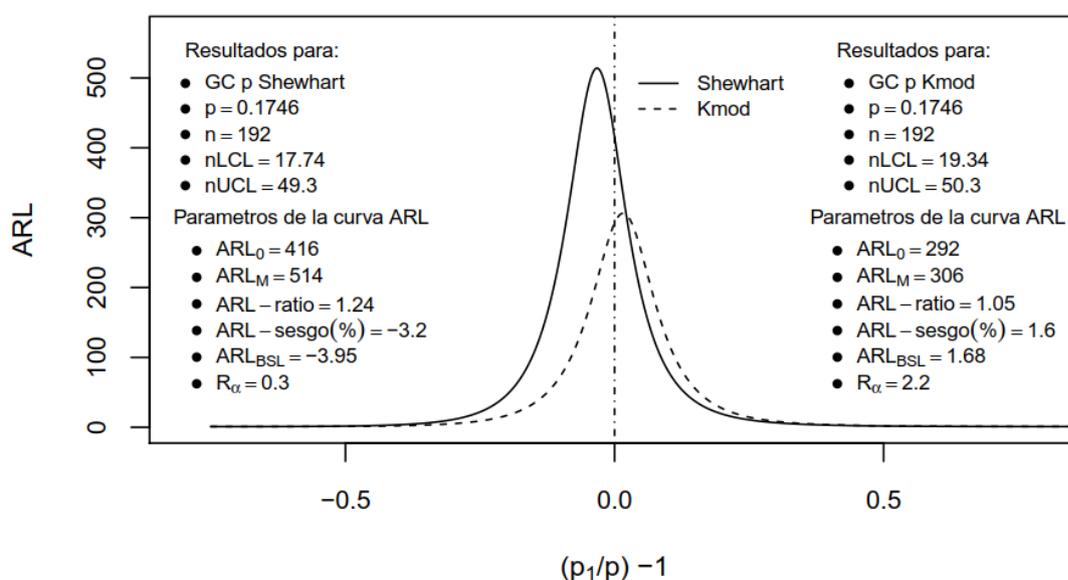


Figura 78. Curvas ARL para los gráficos p Shewhart y Kmod para $p = 0.1746$ y $n = 192$

6 Conclusiones generales, principales aportaciones y posibles temas de investigación

6.1 Conclusiones generales

Las conclusiones generales que se dan en esta sección están relacionadas al logro de los objetivos definidos en la sección 2.3.

Cabe mencionar que aparte de esta sección, en las secciones 4.1.6, 4.2.5 y 5, se dan conclusiones parciales que deben ser tomadas en cuenta a la hora de valorar la profundidad del trabajo realizado y los logros obtenidos de la investigación.

6.1.1 Conclusiones concernientes al objetivo específico número uno

Objetivo específico 1:

Desarrollar parámetros para cuantificar de forma efectiva el sesgo de las curvas ARL para gráficos de control por atributos.

Para lograr este objetivo se concibió el parámetro ARL_{BSL} , el cual se detalla en la sección 5.4, y que sirve para cuantificar la severidad del sesgo en las curvas ARL de gráficos de control por atributos. Mediante este parámetro es posible caracterizar la severidad del sesgo en las curvas ARL de cualquier gráfico de control por atributos del tipo uni-variante (aunque es posible que también pueda ser usado en el contexto multi-variante).

Cabe destacar que el ARL_{BSL} fue de fundamental importancia para poder llegar a cabo la investigación que se resume en esta tesis, siendo de especial relevancia en la aplicación del criterio que dice “*un gráfico de control es considerado como idóneo para monitorizar procesos solo cuando su curva ARL es cuasi insesgada y su ARL_0 es aceptable*” y que sirvió para obtener los resultados más significativos. Cabe destacar que para poder utilizar este criterio se establecieron: i) el rango de ARL_{BSL} para identificar si una curva ARL es o no cuasi insesgada (ver sección 5.4.2) y ii) el rango de ARL_0 aceptables (ver sección 5.1.3).

Mediante el ARL_{BSL} y el ARL_0 , se logró determinar de una forma objetiva la capacidad de monitorización de todos los gráficos de control incluidos en esta tesis, lo que a su vez hizo posible realizar comparaciones entre gráficos existentes y también desarrollar nuevas propuestas de gráficos de control.

6.1.2 Conclusiones concernientes al objetivo específico número dos

Objetivo específico 2:

Desarrollar métodos que faciliten la estimación de la capacidad de monitorización de gráficos de control por atributos.

Para lograr este objetivo se concibieron los siguientes métodos:

1. Para el caso de gráficos de control Shewhart
 - Un método para estimar la severidad del sesgo en las curvas ARL mediante una ecuación simple, ver sección 5.3.1.
 - Un método para estimar la severidad del sesgo en las curvas ARL mediante una técnica gráfica, ver sección 5.3.2.
2. Para el caso de los gráficos de control u y p Kmod
 - Un método basado en el ratio de las colas alfa que sirve para verificar si una curva ARL es o no cuasi insesgada. Este método se detalla en la sección 5.5.3.4 (para el caso del GC p Kmod) y en la sección 5.6.3.4 (para el caso del GC u Kmod).

De los métodos antes citados mención especial se da al desarrollado para los gráficos Kmod por la siguiente razón:

Cuando se construye un gráfico de control por atributos, lo habitual es simplemente asumir que este tendrá una capacidad de monitorización idónea, es decir que podrá detectar tanto mejoras como deterioros del proceso de igual forma. Sin embargo, en esta investigación se ha comprobado que esta presunción no es correcta ya que todos los gráficos por atributos existentes, al menos los incluidos en esta tesis, tienen una capacidad de monitorización que varía en función de los parámetros propios de las distribuciones binomial y de Poisson (es decir p y n para la binomial, u y n para la de Poisson).

Ahora bien, para poder determinar si un gráfico tiene o no una capacidad de monitorización idónea, se deberían computar los parámetros de la curva ARL expuestos en la sección 4.1.8.1, y realizar una evaluación en base a los valores obtenidos. Sin embargo, obtener esos valores no es del todo simple ya que implican realizar cálculos que son largos y complejos, los cuales, sobre todo para personal que no tiene formación en estadística y/o programación algorítmica, serían un poco difíciles de realizar.

Es justamente pensando en esta situación por lo que se concibió el método que acompaña a los gráficos Kmod. Este método permite valorar la capacidad de monitorización de estos gráficos de una manera muy sencilla y además es de fácil aplicación inclusive para personal

que desconoce de estadística. Por otra parte, y desde el punto de vista de logros de la investigación, se considera que este método agrega valor a los gráficos propuestos en esta tesis.

6.1.3 Conclusiones concernientes al objetivo específico número tres

Objetivo específico 3:

Caracterizar la capacidad de monitorización de gráficos de control por atributos relevantes que hayan sido propuestos hasta la fecha de realización de la investigación.

Para lograr este objetivo se caracterizó la capacidad de monitorización de los siguientes gráficos de control:

1. Para el caso de gráficos de control basados en la distribución binomial
 - El gráfico de control p Shewhart, ver sección 5.5.1.
 - Todos los gráficos de control incluidos en la sección 5.5.2 y detallados en la Tabla 12 y en la Tabla 15.
2. Para el caso de gráficos de control basados en la distribución de Poisson
 - El gráfico de control u Shewhart, ver sección 5.6.1.
 - Todos los gráficos de control incluidos en la sección 5.6.2 y detallados en la Tabla 19.

La caracterización de la capacidad de monitorización de los gráficos por atributos antes mencionados, permitió poder compararlos objetivamente, lo que a su vez posibilitó identificar a los que sobresalían. Por ejemplo, en la sección 5.5.1 se identificó que los gráficos basados en “optimización de límites” eran superiores a los basados en “aproximaciones no lineales”, algo que solo fue posible afirmar luego de haber hecho el estudio de caracterización.

Por otra parte, los resultados de las caracterizaciones también posibilitaron establecer que los gráficos por atributos existentes no eran totalmente idóneos, algo de crucial importancia debido a que esa conclusión permitió determinar que si era factible crear nuevos gráficos que fuesen superiores a los ya existentes. Los resultados también ayudaron a identificar las técnicas matemáticas que podrían ser usadas para desarrollar esos nuevos gráficos y los valores ARL a igualar o superar.

6.1.4 Conclusiones concernientes al objetivo específico número cuatro

Objetivo específico 4:

Desarrollar nuevas propuestas de gráficos de control por atributos.

Para lograr este objetivo se desarrollaron los siguientes gráficos de control:

1. El gráfico de control p Kmod
Este gráfico se detalla en la sección 5.5.3.
2. El gráfico de control u Kmod
Este gráfico se detalla en la sección 5.6.3.

Sin ánimo de redundar en lo ya dicho acerca de estos gráficos, solo cabe mencionar que estos tienen un rendimiento ARL superior a cualquiera de los gráficos (p y u respectivamente) incluidos en esta tesis, y que además cuentan con un novedoso método de fácil aplicación que sirve para determinar si su capacidad de monitorización es o no idónea.

6.1.5 Conclusiones concernientes al objetivo específico número cinco

Objetivo específico 5:

Publicar los resultados de la investigación en actas de congresos y en revistas científicas.

Las publicaciones obtenidas de la investigación son las siguientes:

1. Publicaciones en actas de congresos

Argoti, M. A., & Carrión García, A. (2017). A simplified method to calculate the relative ARL-bias in Shewhart p -Charts. *Proceedings of the 8th DQM International Conference* (pp. 239-246). Prijevor - Serbia: DQM Research Center.

Argoti, M. A., Carrión-García, A., & Mulema, A. (2018b). On nonlinear transformation methods used for fraction non conforming process monitoring. *Proceedings of the 9th DQM International Conference* (pp. 26-38). Prijevor - Serbia: DQM Research Centre.

2. Artículos publicados en revistas internacionales

Argoti, M. A., & Carrión-García, A. (2018a). A novel approach for estimating the ARL-bias severity of Shewhart p -Charts. *International Journal for Quality Research*, 12(1), 209-226 , DOI: 10.18421/IJQR12.01-12.

Argoti, M. A., & Carrión-García, A. (2019a). A heuristic method for obtaining quasi ARL-unbiased p -Charts. *Quality and Reliability Engineering International*, 35(1), 47-61. DOI: 10.1002/qre.2379.

3. Artículos actualmente bajo revisión en revistas internacionales

Argoti, M. A., & Carrión-García, A. (2019b). A quasi ARL- unbiased u control chart. En revisión.

Sin lugar a duda el logro más satisfactorio del trabajo de investigación resumido en esta tesis, es haber conseguido que sus resultados más significativos hayan sido publicados. En este aspecto mención especial se da al artículo “A heuristic method for obtaining quasi ARL-unbiased p-Charts” que fue publicado en la revista “Quality and Reliability Engineering International”, siendo esta una revista de alto prestigio en las ramas de ingeniería y estadística industrial y que consta en el listado JCR (Journal Citation Report) con un factor de impacto que al momento es de 1.604.

6.2 Principales aportaciones obtenidas de la investigación

En esta sección se mencionan, sin entrar en gran detalle ni profundidad, las principales aportaciones obtenidas de la investigación, que se consideran son las siguientes:

1. Haber caracterizado el comportamiento de las colas alfa de los gráficos de control por atributos basados en las distribuciones binomial y de Poisson, y determinado que su valor, en vez de ser constante, más bien varía y fluctúa en función de los parámetros propios de esas distribuciones (es decir p y n para la binomial, u y n para la de Poisson). Las secciones relevantes son: 4.1.6, 4.2.5, 5.5.2 y 5.6.2.
2. Haber caracterizado el comportamiento del ARL_0 en los gráficos de control por atributos basados en las distribuciones binomial y de Poisson y determinado que su valor no es constante, si no que más bien este varía y fluctúa en función de los parámetros propios de esas distribuciones (es decir p y n para la binomial, u y n para la de Poisson). Las secciones relevantes son: 4.1.6, 4.2.5, 5.5.2 y 5.6.2.
3. Haber desarrollado una propuesta de rango de valores ARL_0 aceptables para gráficos de control por atributos, ver sección 5.1.3.
4. Haber desarrollado el parámetro R_α (o ratio de las colas alfa), ver sección 5.1.4.
5. Haber desarrollado dos métodos simples que sirven para estimar la severidad del sesgo de las curvas ARL para gráficos de control p Shewhart, ver sección 5.3.
6. Haber desarrollado el parámetro ARL-ratio (ARL_0/ARL_M) que sirve para valorar el sesgo de las curvas ARL, ver sección 5.4.
7. Haber desarrollado en parámetro ARL_{BSL} para cuantificar la severidad del sesgo de las curvas ARL, ver sección 5.4.

8. Haber desarrollado el concepto y definición de *curvas ARL cuasi insesgadas*, además del criterio ARL_{BSL} para establecer cuando estas se producen, ver secciones 5.3.1.1 y 5.4.
9. Haber determinado que los gráficos de control p y u Shewhart tienen curvas ARL con sesgo negativo, ver secciones 5.5.1. y 5.6.1.
10. Haber determinado que los gráficos de control p y u “alternativos” no son del todo idóneos (o que no siempre tienen una capacidad de monitorización ideal). ver secciones 5.5.2 y 5.6.2.
11. Haber desarrollado dos nuevos gráficos de control, el p Kmod (ver sección 5.5.3) y el u Kmod (ver sección 5.6), que son superiores a cualquier otro gráfico incluido en esta tesis.
12. Haber desarrollado un método basado en el ratio de las colas alfa que sirve para determinar, de una manera muy fácil, si las curvas ARL de los gráficos Kmod son o no cuasi insesgadas.

6.3 Posibles temas para futura investigación

Algunos posibles temas para futura investigación son:

1. Desarrollar un gráfico de control CUSUM p Kmod
2. Desarrollar un gráfico de control EMWA p Kmod
3. Desarrollar un gráfico de control CUSUM u Kmod
4. Desarrollar un gráfico de control EMWA u Kmod
5. Aplicar a los gráficos Kmod distintas estrategias de muestreo (VSS, VSI, ...etc)

7 Referencias

- Acosta-Mejía, C. (1999). Improved p-charts to monitor process quality. *IEE Transactions*, 31(6), 509-516.
- Agresti, A., & Brian, C. (2000). Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Results from Adding Two Success and Two Failures. *The American Statistician*, 54(4), 280-288.
- Agresti, A., & Coull, B. A. (1998). Aproximate is better than "Exact" for interval Estimation of binomial proportions. *The American Statistician*, 52(2), 119-126.
- Amin, S. G. (2001). Control Charts 101: A Guide to Health Care Applications. *Quality Management in Healthcare*, 9(3), 1-27.
- Argoti, M. A., & Carrión García, A. (2017). A simplified method to calculate the relative ARL-bias in Shewhart p-Charts. *Proceedings of the 8th DQM International Conference* (pp. 239-246). Prijevor - Serbia: DQM Research Center.
- Argoti, M. A., & Carrión-García, A. (2018a). A novel approach for estimating the ARL-bias severity of Shewhart p-Charts. *International Journal for Quality Research*, 12(1), 209-226 , DOI: 10.18421/IJQR12.01-12.
- Argoti, M. A., & Carrión-García, A. (2019a). A heuristic method for obtaining quasi ARL-unbiased p-Charts. *Quality and Reliability Engineering International*, 35(1), 47-61. DOI: 10.1002/qre.2379.
- Argoti, M. A., & Carrión-García, A. (2019b). A quasi ARL- unbiased u control chart. En revisión.
- Argoti, M. A., Carrión-García, A., & Mulema, A. (2018b). On nonlinear transformation methods used for fraction non conforming process monitoring. *Proceedings of the 9th DQM International Conference* (pp. 26-38). Prijevor - Serbia: DQM Research Centre.
- Bersimis, S., Psarakis, S., & Panaretos, J. (2006). Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International*, 23(5), 517-543.
- Brown, L. D., Cai, T. T., & Anirban, D. (2002). Confidence intervals for a binomial proportion and asymptotic expansions. *The annals of statistics*, 30(1), 160-201.

- Brown, L. D., Cai, T. T., & Dasgupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science*, 16(2), 128-133.
- Carey, R. G. (2003). *Improving Healthcare with Control Charts*. Milwaukee, Wisconsin: ASQ Quality Press.
- Cavalcanti, A. B., & Cordeiro, G. M. (2006). An improved u chart for attributes. *Brazilian Journal of Probability and Statistics*, 20, 133-140.
- Champ, C. W., & Woodhall, W. H. (1987). Exact Results for Shewhart Control Charts with Supplementary Runs Rules. *Technometrics*, 29(4), 393-399.
- Chang, C.-H., Lin, J.-J., Pal, N., & Chiang, M.-C. (2008). A Note on Improved Approximation of the Binomial Distribution by the Skew-Normal Distribution. *The American Statistician*, 62(2), 167-170.
- Chen, G. (1998). An improved p chart through simple adjustments. *Journal of Quality Technology*, 30(2), 142-151.
- Chen, G., & Cheng, S. W. (1998). The exact u Chart can be obtained using simple adjustments. *Statistics and Probability Letters*, 37, 357-365.
- Chen, G., Cheng, S. W., & Xie, H. (2004). A new EWMA control chart for monitoring both location and dispersion. *Quality Technology & Quantitative Management*, 2, 217-231.
- Chen, K. S., Chang, T. C., Wang, K. J., & Huang, C. T. (2015). Developing control charts in monitoring service quality based on the number of customer complaints. *Total Quality Management & Business Excellence*, 26(5-6), 675-689.
- Deming, W. E. (1989). *Calidad, Productividad y Competitividad: La salida de la crisis*. Madrid: Diaz Santos, S. A.
- Dhini, A., & Surjandari, I. (2016). Review on Some Multivariate Statistical Process Control. *Proceedings of the 2016 International Conference on Industrial Engineering and Operations Management*. Kuala Lumpur.
- Duncan, A. J. (1986). *Quality Control and industrial Statistics* (Vol. 5th). Homewood: Irwin.
- Duncan, A. J. (2000). *Control de Calidad y Estadística Industrial*. Santafé de Bogotá: ALFAOMEGA.
- Eriksson, L., Byrne, T., Johansson, E., Trygg, J., & Vikström, C. (2013). *Multi- and Megavariate Data Analysis: Basic Principles and Applications* (3ra ed.). MKS Umetrics AB.

-
- Fuchs, C., & Kenett, R. S. (1998). *Multivariate Quality Control: Theory and Applications*. New York: Chapman and Hall/CRC.
- Gan, F. F. (1993). An Optimal Design of CUSUM Control Charts for Binomial Counts. *Journal of Applied Statistics*, 20, 445-460.
- Garvin, D. A. (1984). What does "Product Quality Really Mean? *Sloan Management Review*, 25-43.
- Garvin, D. A. (1987). Competing in the Eight Dimensions of Quality. *Havard Bussiness Review*, 101-109.
- Ishikawa, K. (1994). *Introducción al control de la Calidad*. Madrid: Diaz de Santos.
- Joekes, S., & Pimentel-Barbosa, E. (2013). An improved attribute control chart for monitoring non-conforming proportion in high quality processes. *Control Engineering Practice*, 21(4), 407-412 , DOI: 10.1016/j.conengprac.2012.12.005.
- Juran, J. M. (1990). *Juran y la planificación para la calidad*. Madrid: Diaz de Santos, S.A.
- Khoo, M. B. (2003). Increasing the Sensitivity of Control Chart for Fraction Nonconforming. *Quality Engineering*, 16(2), 309-319.
- Kittlitz, R. G. (2006). Calculating the (almost) exact control limits for a C-chart. *Quality Engineering*, 18(3), 356-366.
- Kourti, T., & MacGregor, J. F. (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, 28, 409-428.
- Lowry, C. A., & Montgomery, D. C. (1995). A review of multivariate control charts. *IIE Transactions*(6), 800-810.
- Lucas, J. M. (1985). Counted Data CUSUM's. *Technometrics*, 27(3), 129-144.
- Mengesha, A., Singh, A. P., & Amedie, W. A. (2013). Quality improvement using statistical process control tools in glass bottles manufacturing company. *International Journal for Quality Reseach*, 7(1), 107-126.
- Miranda Gonzáles, F. J., Lacoba, S. R., & Chamorro Mera, A. (2007). *Introcucion a la Gestion de la Calidad*. Madrid: Delta Publicaciones Universitarias.
- Mitra, A. (2008). *Fundamentals of Quality Control and Improvement* (3th ed.). Hoboken, Canada: Jhon Wiley & Sons.

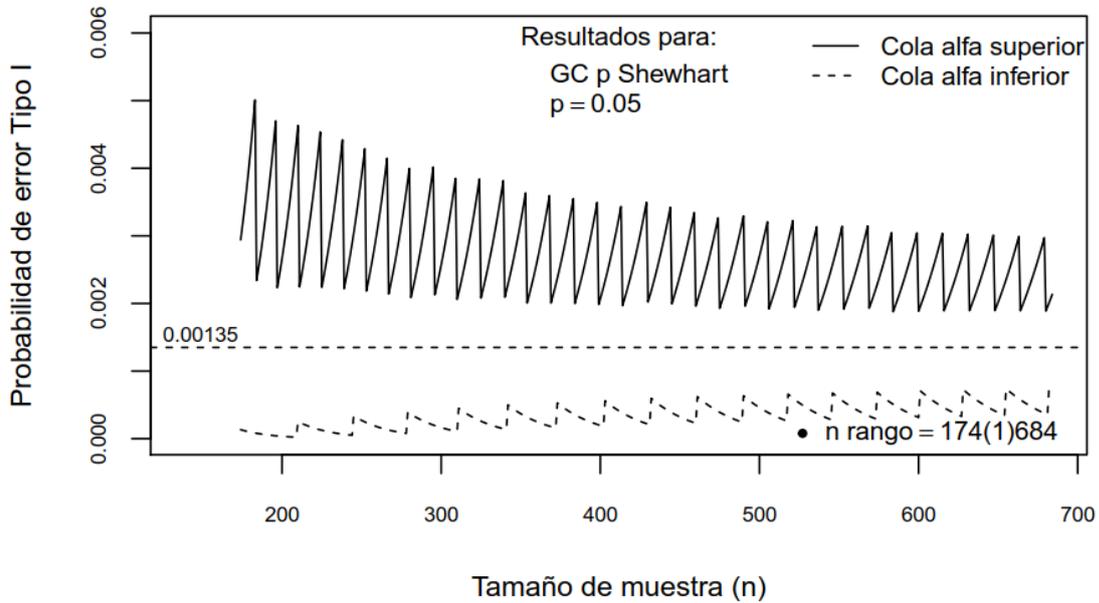
-
- Montgomery, D. C. (2009). *Introduction to Statistical Quality Control* (6th ed.). Jefferson City, United States of America: Jhon Wiley & Sons, Inc.
- Morais, M. C. (2016). An ARL-Unbiased np-Chart. *Economic Quality Control*, 31(1), 11-21.
- Moreno-Luzón, M. D., Peris, F. J., & González, T. F. (2001). *Gestion de la Calidad y Diseño de Organizaciones: Teoría y estudio de casos*. Prentice Hall.
- Morton, A., Whitby, M., McLaws, M. L., & Sartor, A. (2002). The application of statistical process control charts to the detection and monitoring of hospital-acquired infections. *21*(4), 112-117.
- Nelson, L. S. (1984). The Shewhart Control Chart—Tests for Special Causes. *Journal of Quality*, 16(4), 88-89.
- Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 41, 100-115.
- Park, C. (2013). An Improved p Chart Based on the Wilson Interval. *Journal of Statistics and Management Systems*, 16:2-3, 201-221.
- Qiu, P. (2014). *Introduction to Statistical Process Control*. Boca Raton, FL: CRC Press.
- Quesenberry, C. P. (1991). SPC Q Charts for a Binomial Parameter: Short and Long Runs. *Journal of Quality Technology*, 23, 239-246.
- Quesenberry, C. P. (1991-2). SPC Q charts por a Poisson parameter: Short and long runs. *Journal of Quality and Technology*, 23(4), 293-303.
- Quesenberry, C. P. (1995). On properties of binomial Q charts for attributes. *Journal of Quality Technology*, 27, 204-213.
- Rich, W. J. (2009). "Examining the accuracy of the normal distribution to the poisson random variable. *Master Theses and Doctoral Dissertations*. 262 , <http://commons.emich.edu/theses/262>.
- Roberts, S. W. (1959). Control Chart tests based on geometric moving average. *Technometrics*, 1(3), 239-250.
- Ryan, T. P. (1989). *Statistical Methods for Quality Improvement*. New York: Jhon Wiley & Sons.
- Ryan, T. P. (2011). *Statistical Methods for Quality Improvement* (3th ed.). New York: John Wiley & Sons.

-
- Ryan, T. P., & Schwertman, N. C. (1997). Optimal Limits for attribute control charts. *Journal of Quality Technology*, 29, 86-98.
- Schader, M., & Schmid, F. (1989). Two Rules of Thumb for the Approximation of the Binomial Distribution by the Normal Distribution. *The American Statistician*, 43(1), 23-24.
- Shewhart, W. A. (1939). *Statistical Method from the view point of quality control*. Washington: The Graduate School, The department of agriculture.
- Simons, G., & Johnson, N. L. (1971). On the convergence of binomial to Poisson distributions. *The annals of Mathematical Science*, 5, 1735-1736.
- Topalidou, E., & Psakaris, S. (2009). Review of Multinomial and Multi attribute Quality Control Chart. *Quality and Reliability Engineering International*, 25, 773-804.
- Treanich, A., & Bourke, P. (1993). EWMA Control Charts Using Attributes Data. *Journal of the Royal Statistical Society Series D (The Statistician)*, 42(3), 215.
- Tsai, T. R., Lin, C. C., & Wu, S. J. (2006). Alternative Attribute Control Charts Based on Improved Square Root Transformation. *Tamsui Oxford Journal of Mathematical Sciences*, 22(1), 61-72.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2007). *Probabilidad y estadística para ingeniería y ciencias*. México: Pearson Educación.
- Winterbottom, A. (1993). Simple Adjustments To Improve Control Limits On Attribute Charts. *Quality and Reliability Engineering International*, 9, 105-109.
- Wood, M. (1994). Statistical Methods for Monitoring Service Processes. *International Journal of Service Industry Management*, 5(4), 53-68.
- Yeh, A., Lin, D. K., & McGrath, R. N. (2006). Multivariate Control Charts for Monitoring Covariance Matrix: A Review. *Quality Technology and Qualitative Management*, 3(4), 415-336.
- Zhang, W., & Yu, T. (2005). Weighted-loss-function CUSUM chart for monitoring mean and variance of a production process. *International Journal of Production Research*, 43(14), 3027-3044.

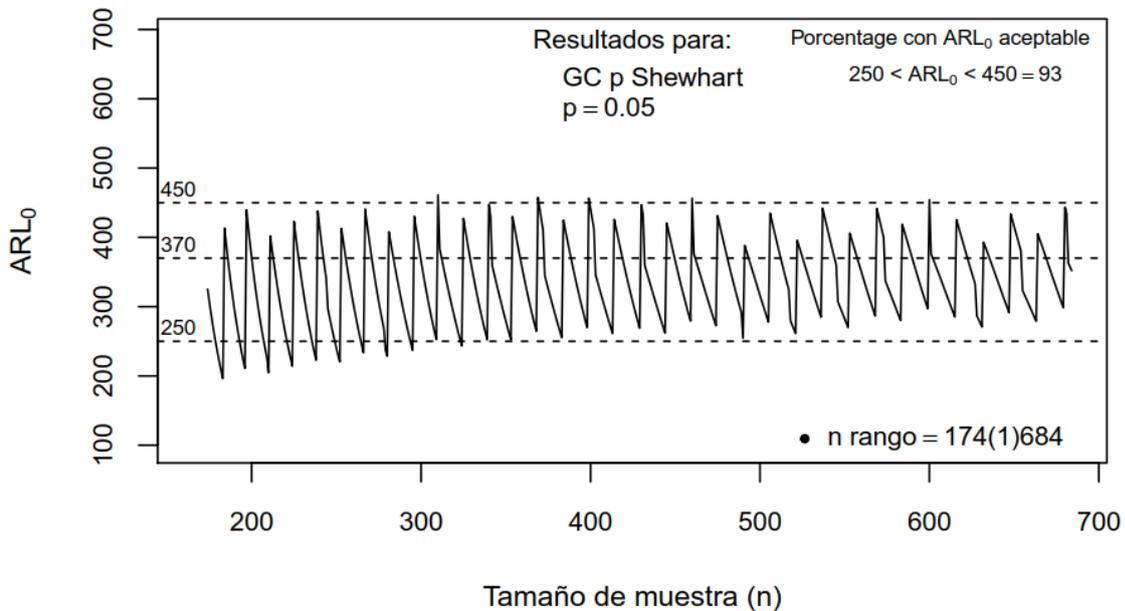
8 Anexos

8.1 Anexo 1

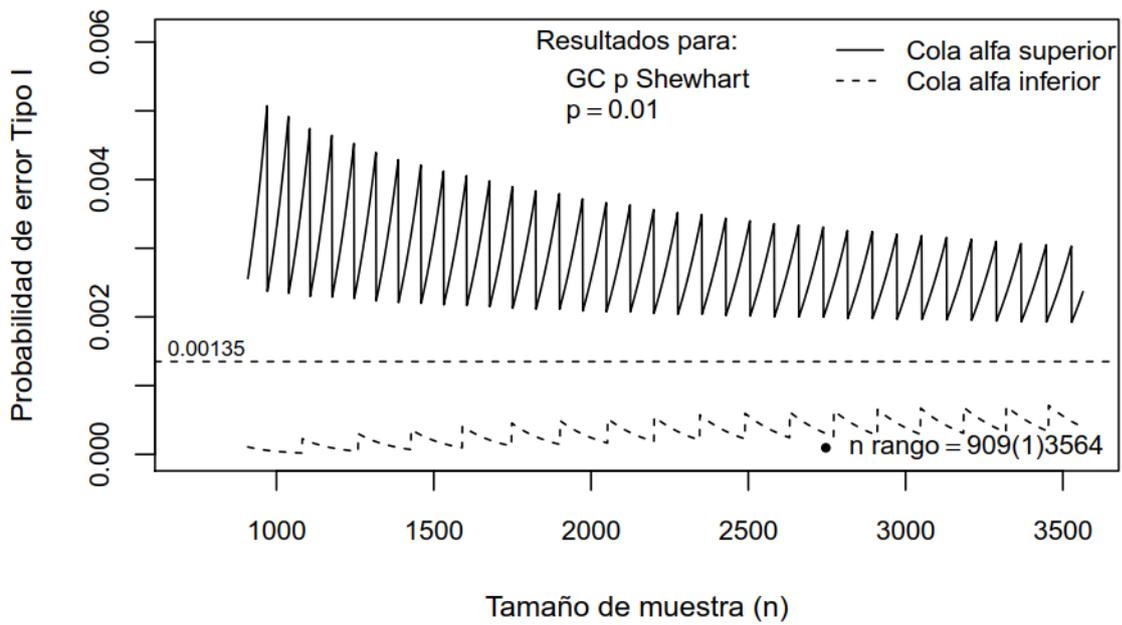
8.1.1 Gráfico de control p Shewhart: Comportamiento de las colas alfa y el ARL_0 (ver sección 4.1.6)



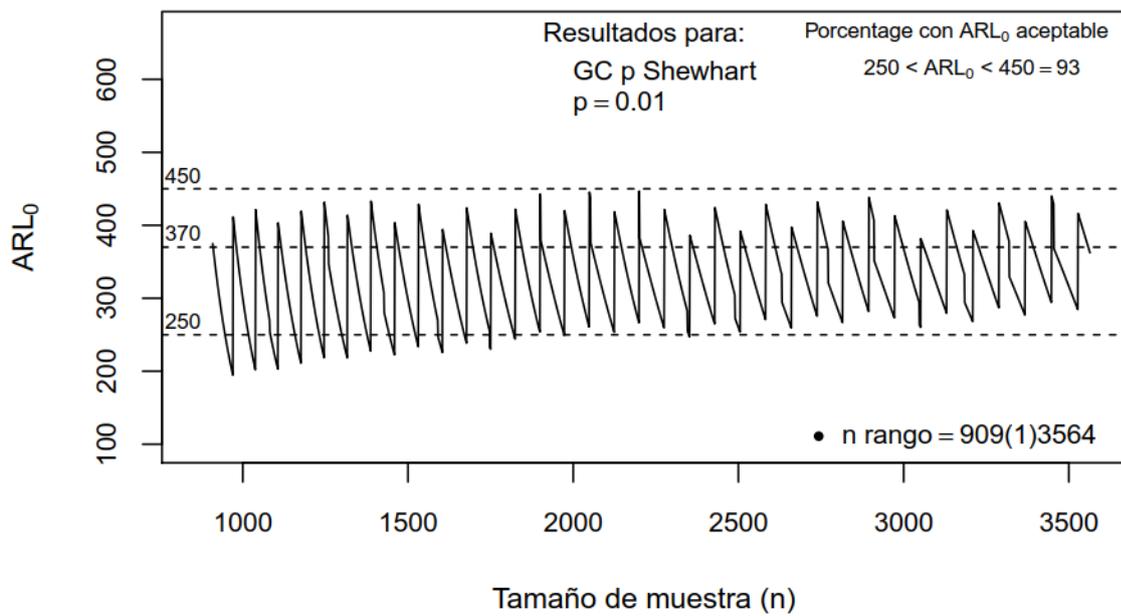
GC p Shewhart - Comportamiento real de la colas alfa para $p = 0.05$



GC p Shewhart – Fluctuación del ARL_0 con $p = 0.05$



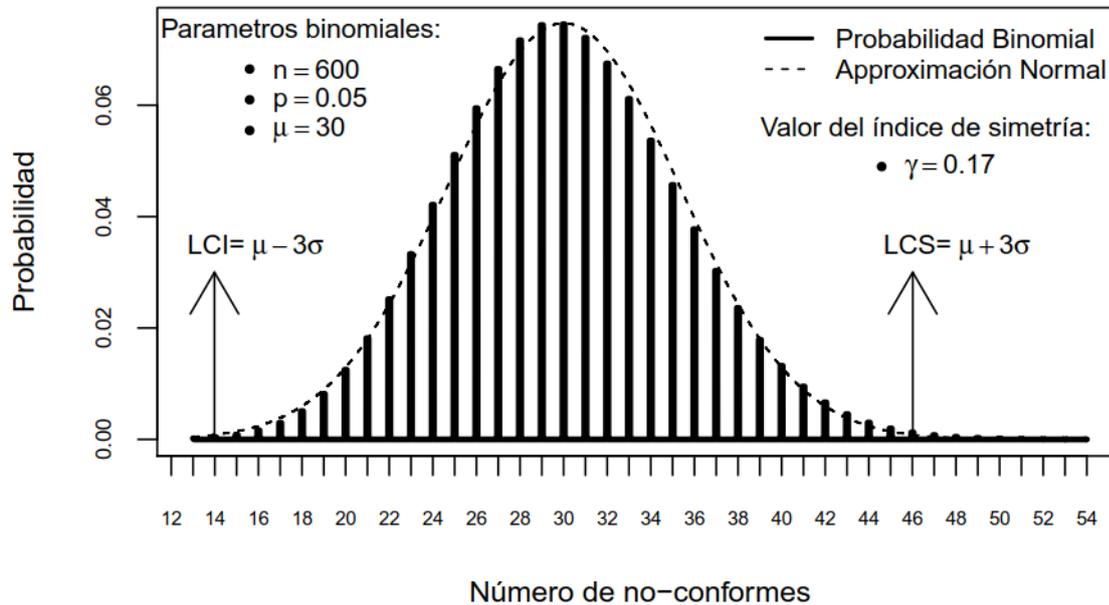
GC p Shewhart - Comportamiento real de la colas alfa para $p = 0.01$



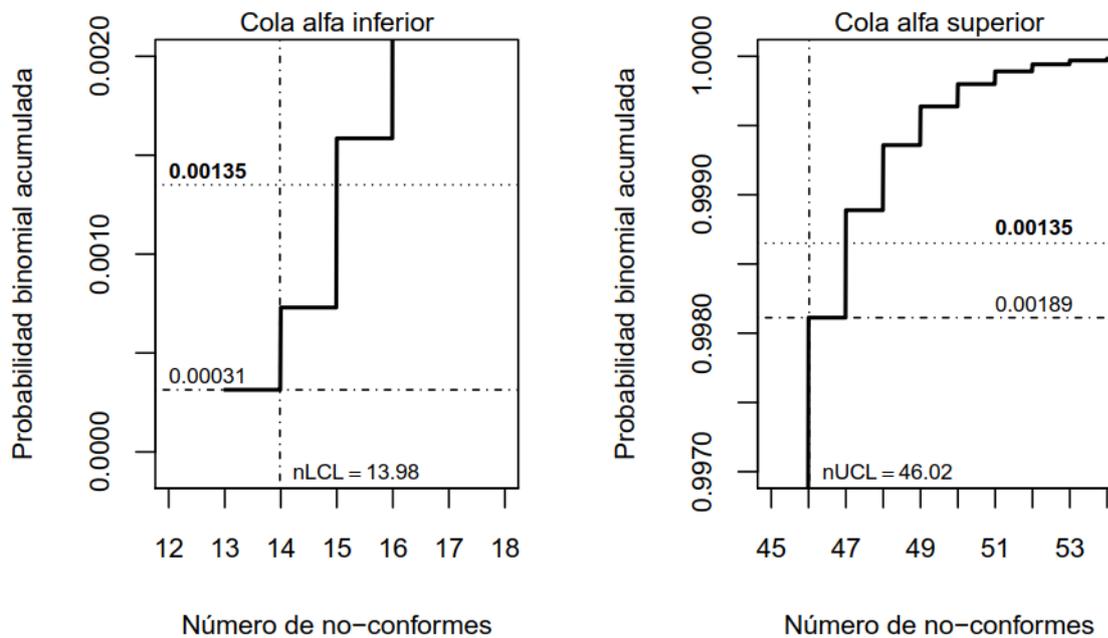
GC p Shewhart – Fluctuación del ARL_0 con $p = 0.01$

8.2 Anexo 2

8.2.1 Ejemplo de ajuste de la aproximación normal a la binomial para $n = 600$ y $p = 0.05$ (ver sección 4.1.7)



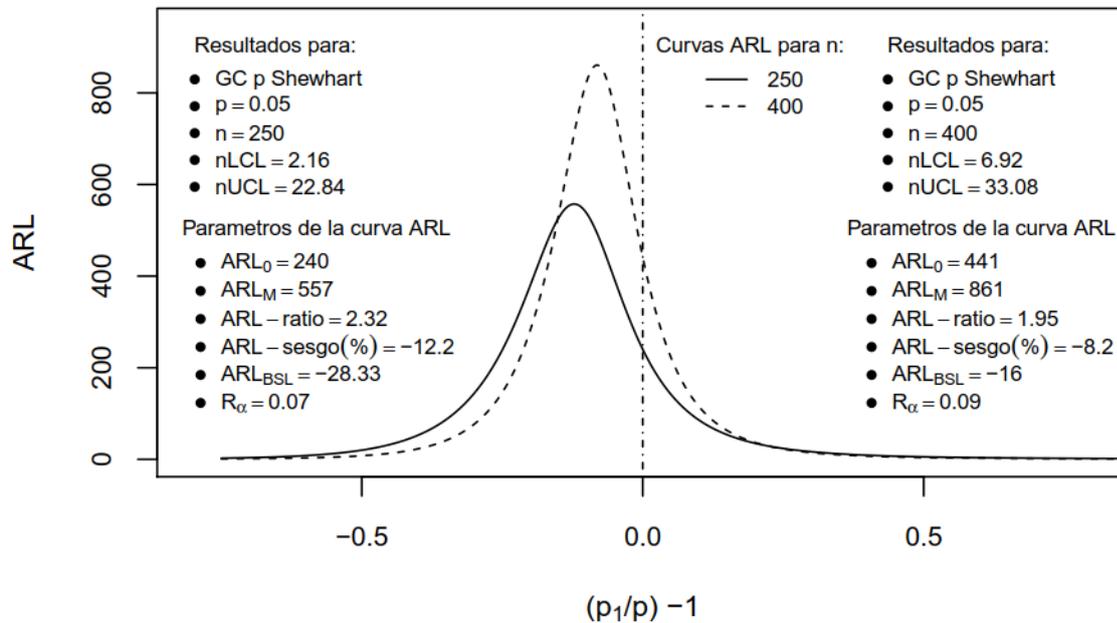
Ajuste de la aproximación normal a la binomial para $n = 600$ y $p = 0.05$



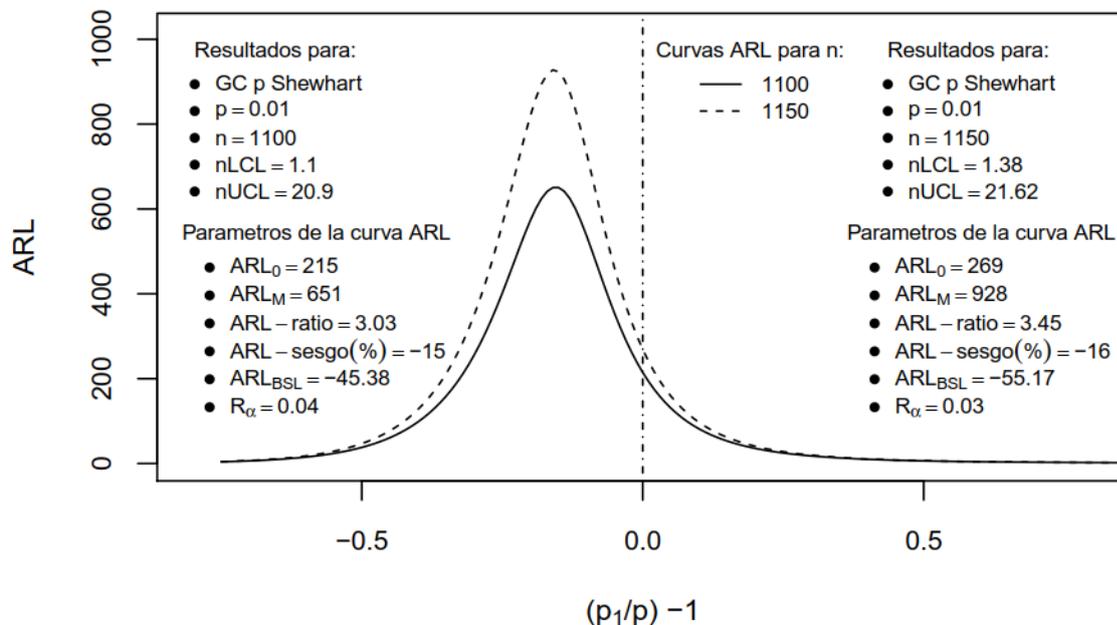
Colas de la probabilidad binomial acumulada para $n = 600$ y $p = 0.05$

8.3 Anexo 3

8.3.1 Gráfico de control p Shewhart: Ejemplos de curvas ARL (ver sección 4.1.8)



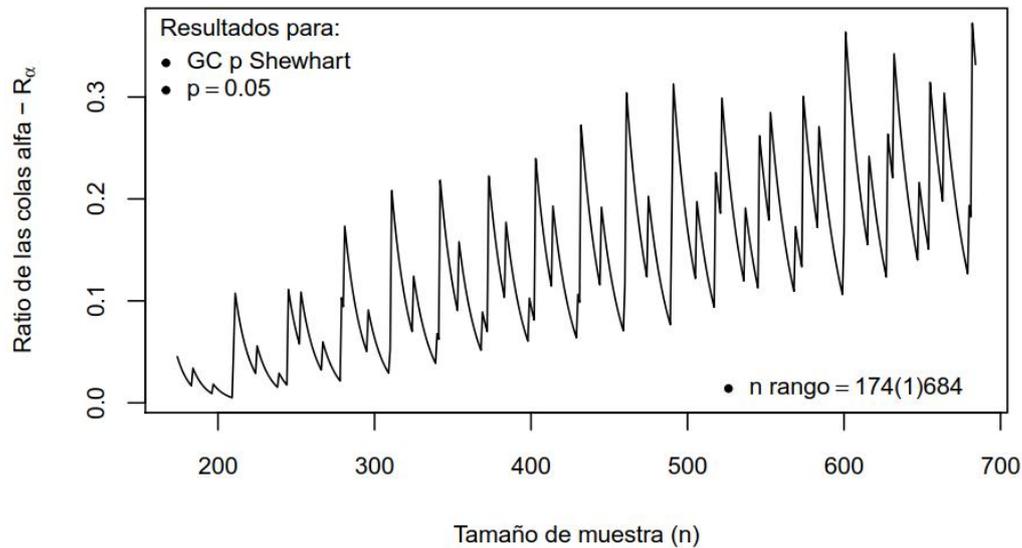
GC p Shewhart – Curvas ARL para $p = 0.05$ y $n = 250$ y 400



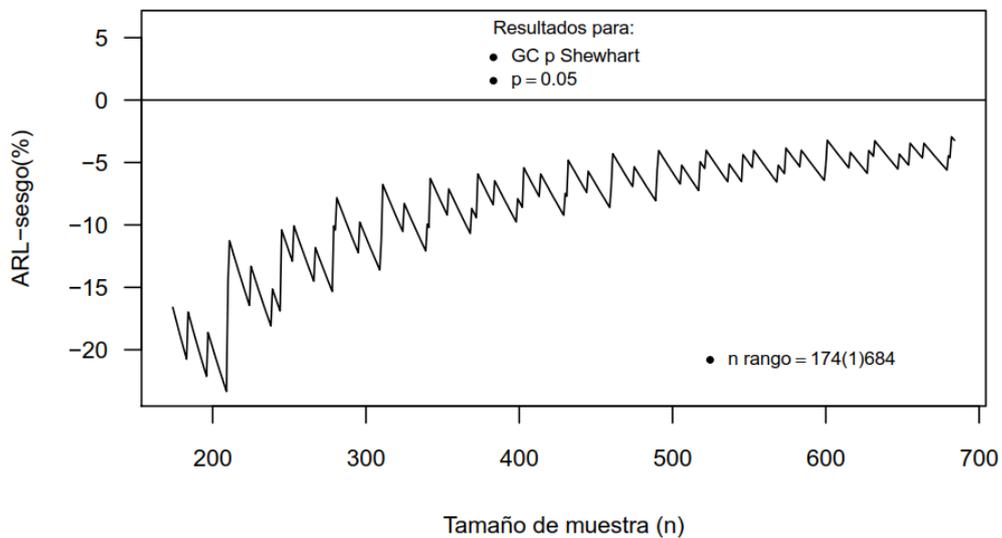
GC p Shewhart – Curvas ARL para $p = 0.01$ y $n = 1100$ y 1150

8.4 Anexo 4

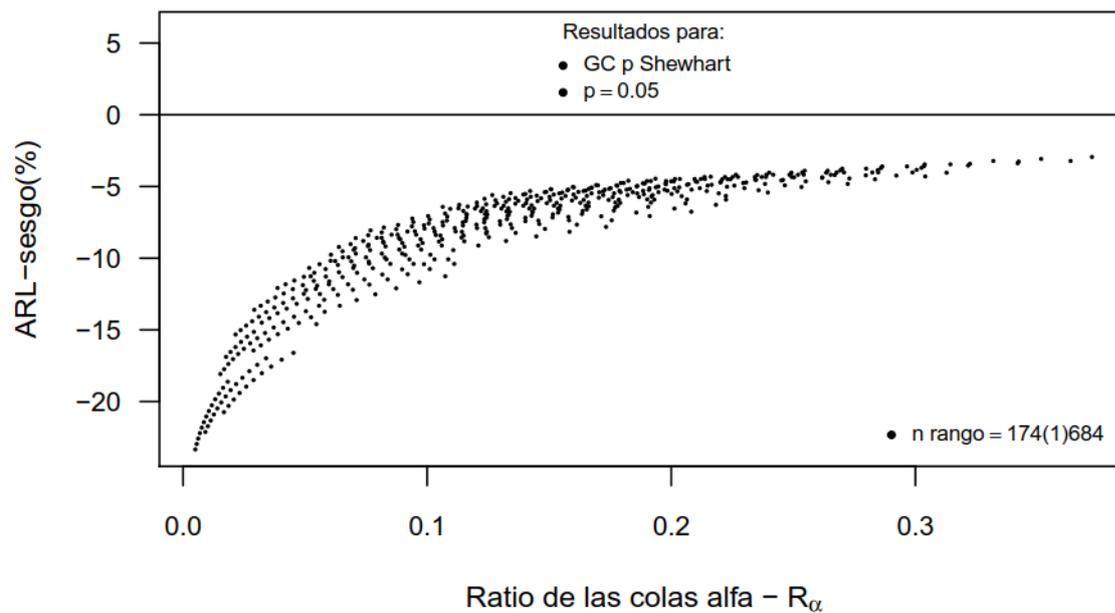
8.4.1 Gráfico de control p Shewhart: Caracterización del ARL-sesgo(%) y R_α (ver sección 5.2)



GC p Shewhart: Variación del R_α para $p = 0.05$



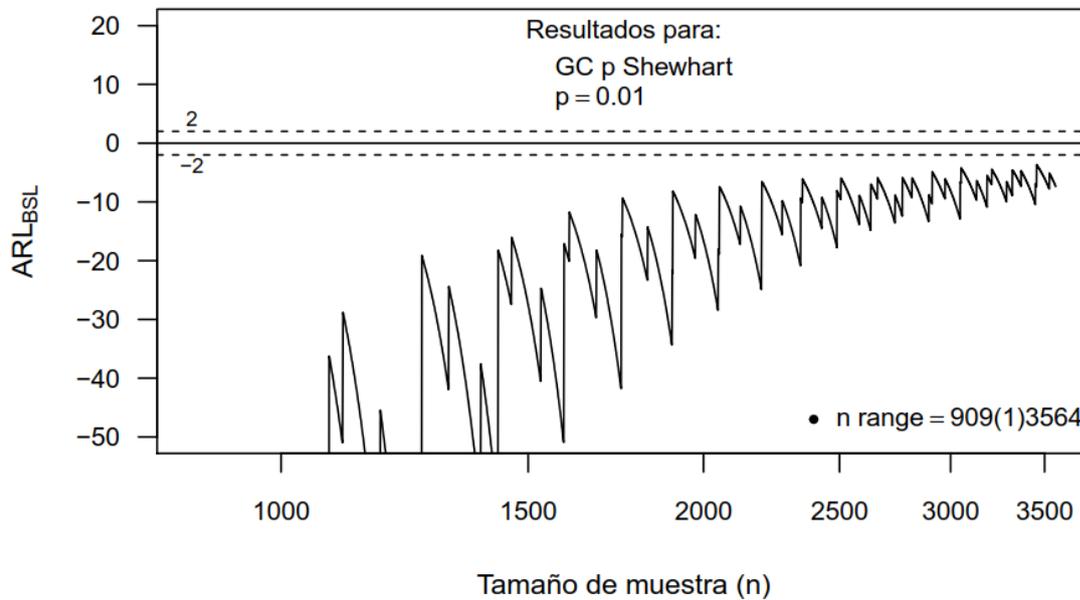
GC p Shewhart: Variación del ARL-sesgo(%) $p = 0.05$



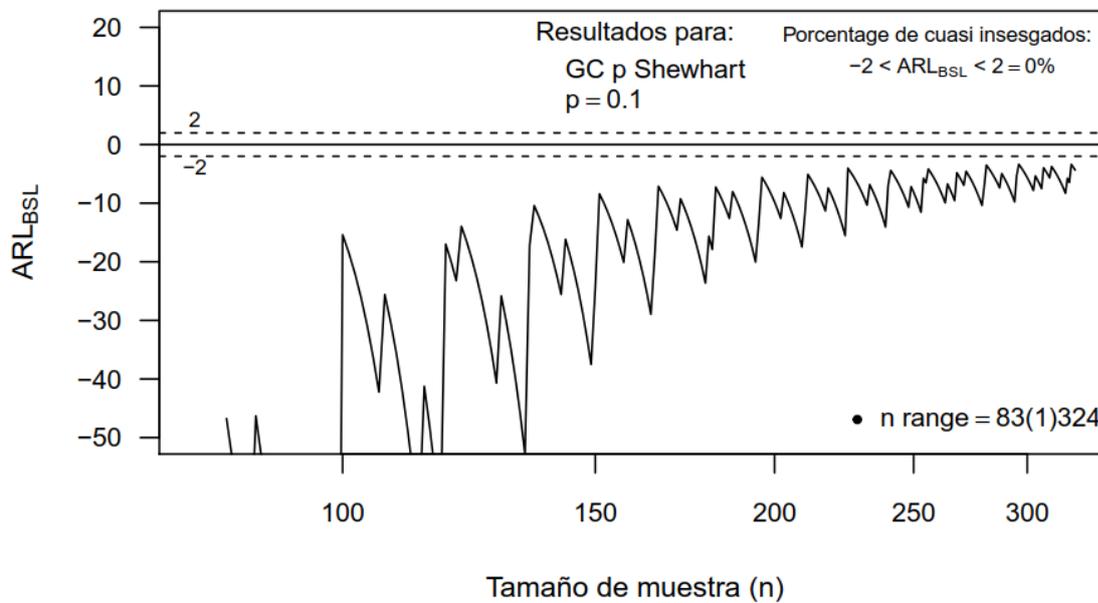
GC p Shewhart: Relación entre ARL-sesgo(%) y R_α para $p=0.05$

8.5 Anexo 5

8.5.1 Gráfico de control p Shewhart: Ejemplos de resultados ARL_{BSL} (ver sección 5.5.1)



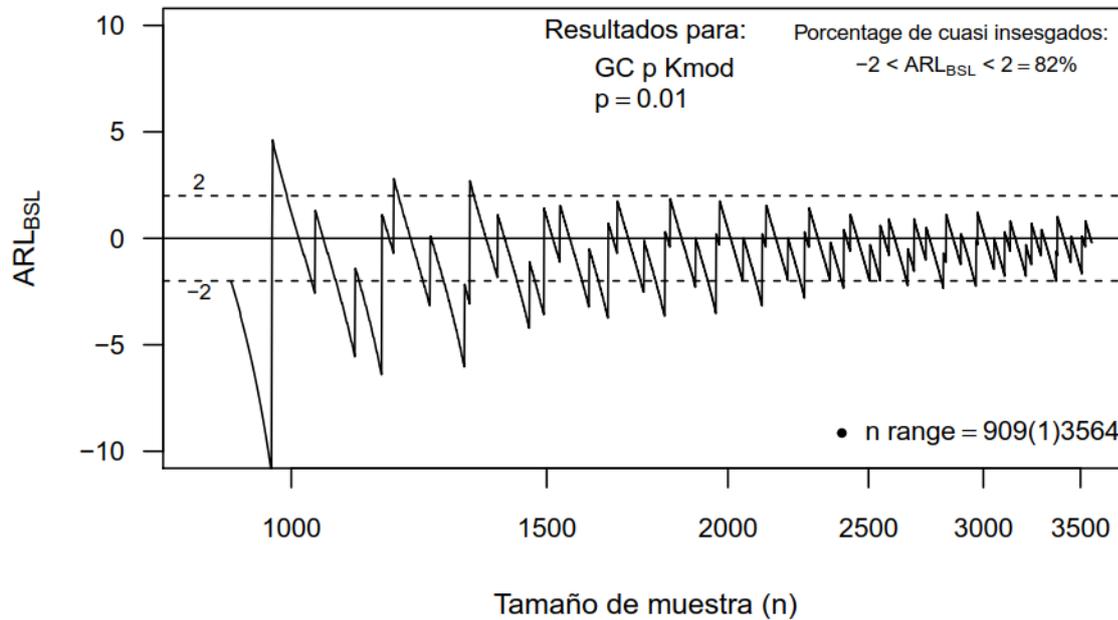
GC p Shewhart: ARL_{BSL} para $p = 0.01$



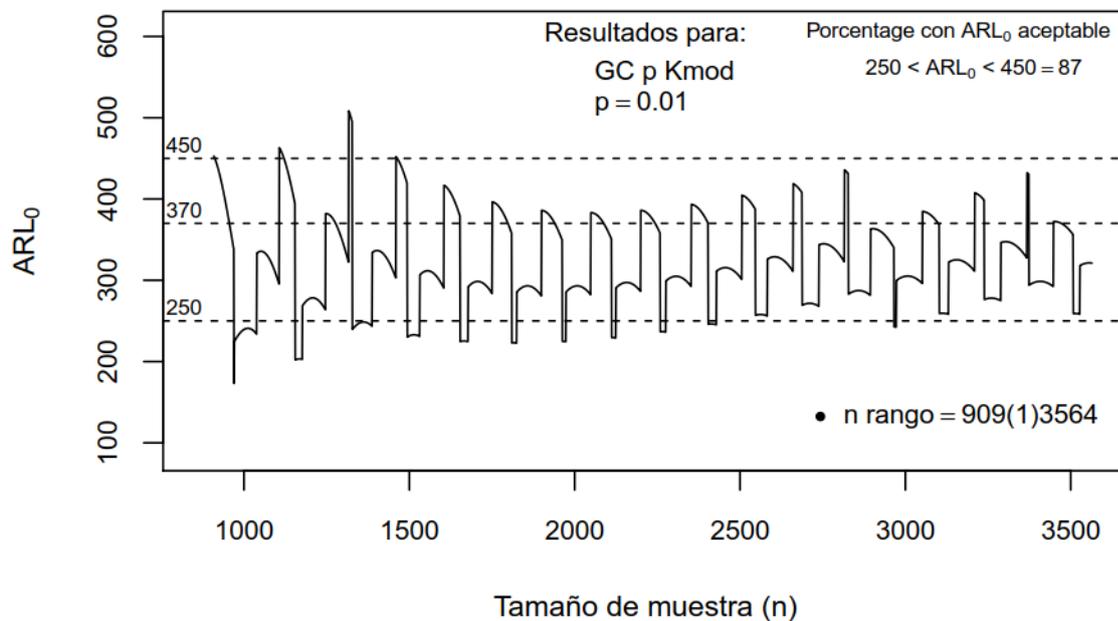
GC p Shewhart: ARL_{BSL} para $p = 0.1$

8.6 Anexo 6

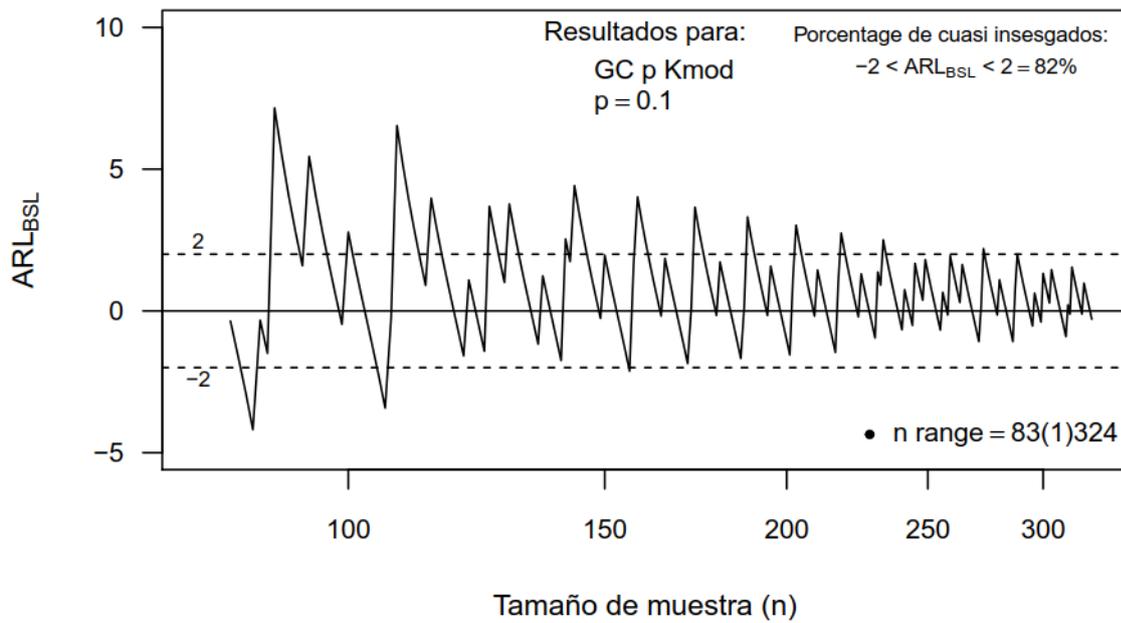
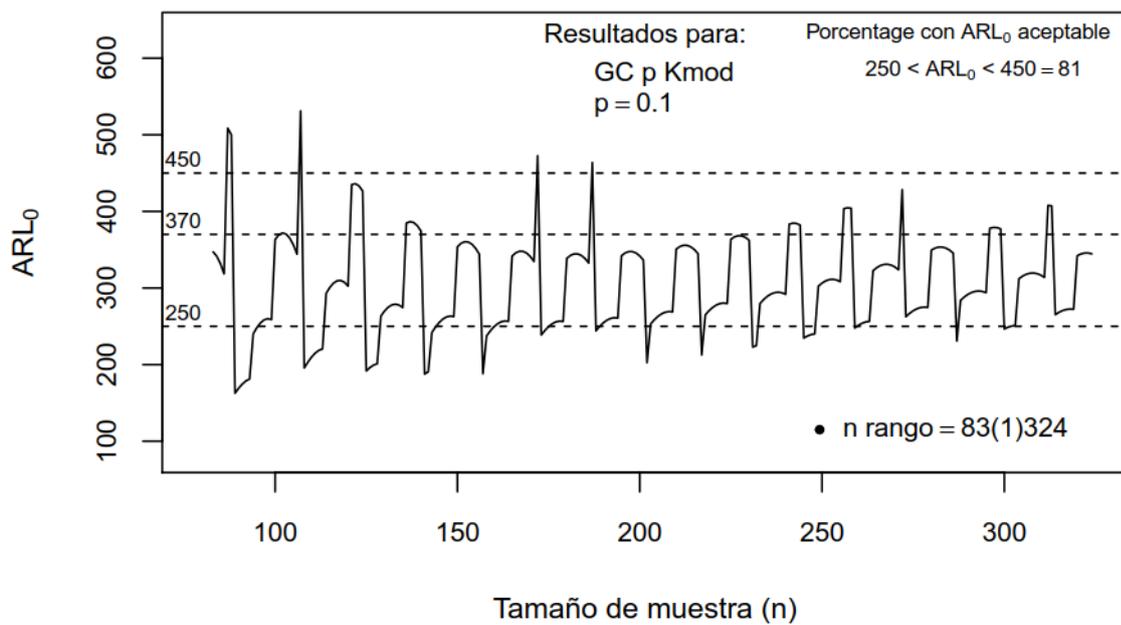
8.6.1 Gráfico de control p K_{mod} : Ejemplos de resultados ARL_{BSL} y ARL_0 (ver sección 5.5.3.2)



GC p K_{mod} : Comportamiento del ARL_{BSL} para $p = 0.01$

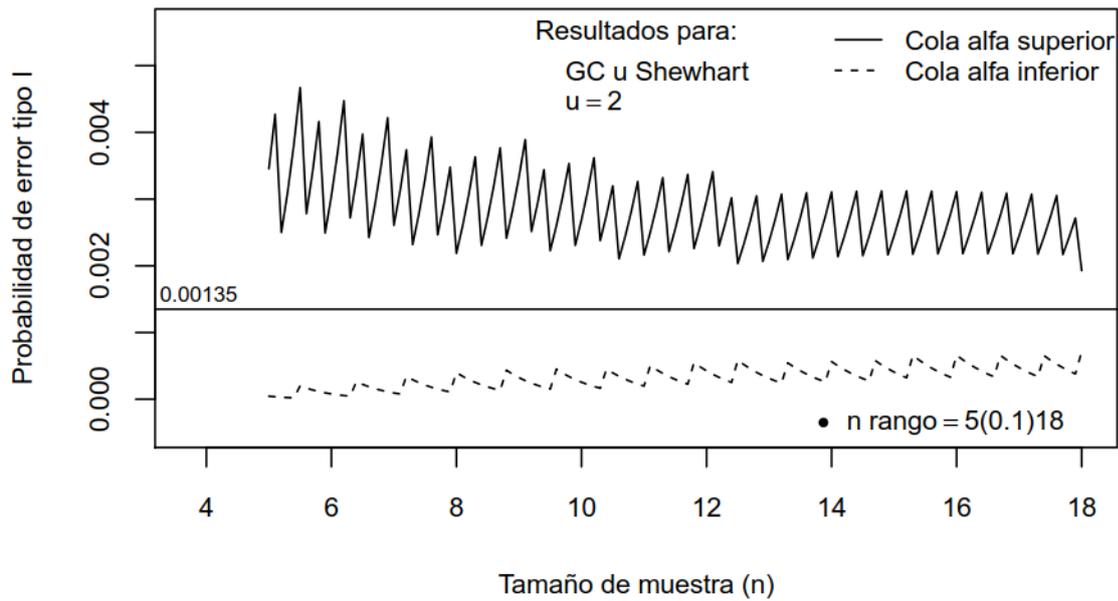


GC p K_{mod} : Comportamiento del ARL_0 para $p = 0.01$

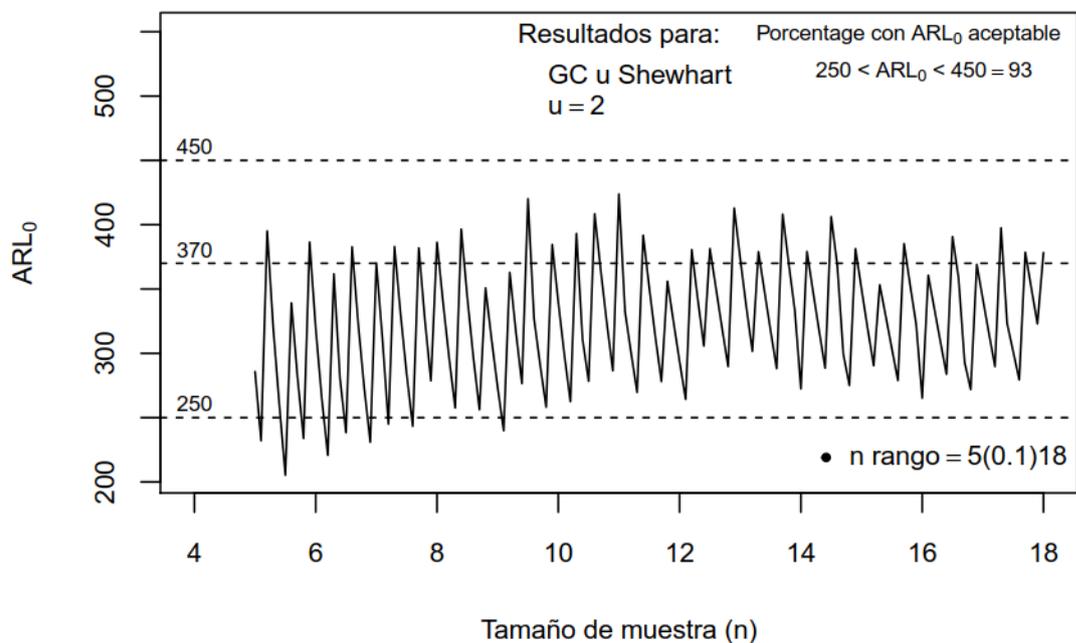
GC p Kmod: Comportamiento del ARL_{BSL} para $p = 0.1$ GC p Kmod: Comportamiento del ARL₀ para $p = 0.1$

8.7 Anexo 7

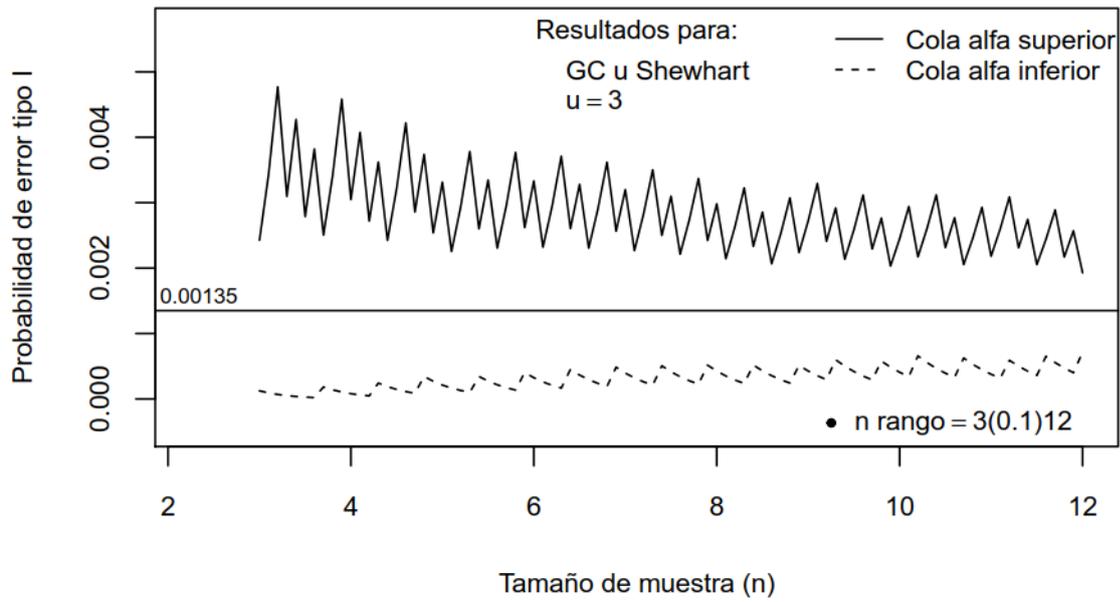
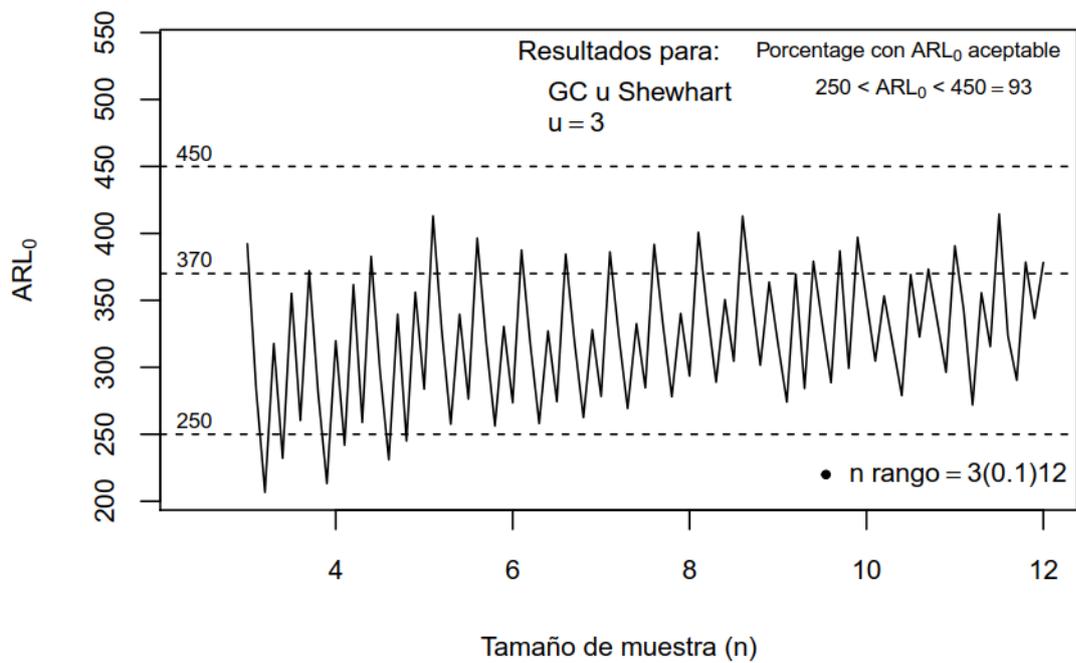
8.7.1 Gráfico de control u Shewhart: Comportamiento de las colas alfa y el ARL_0 (ver secciones 4.2.5.1 y 4.2.5.2)



GC u Shewhart - Comportamiento real de la colas alfa para $u = 2$

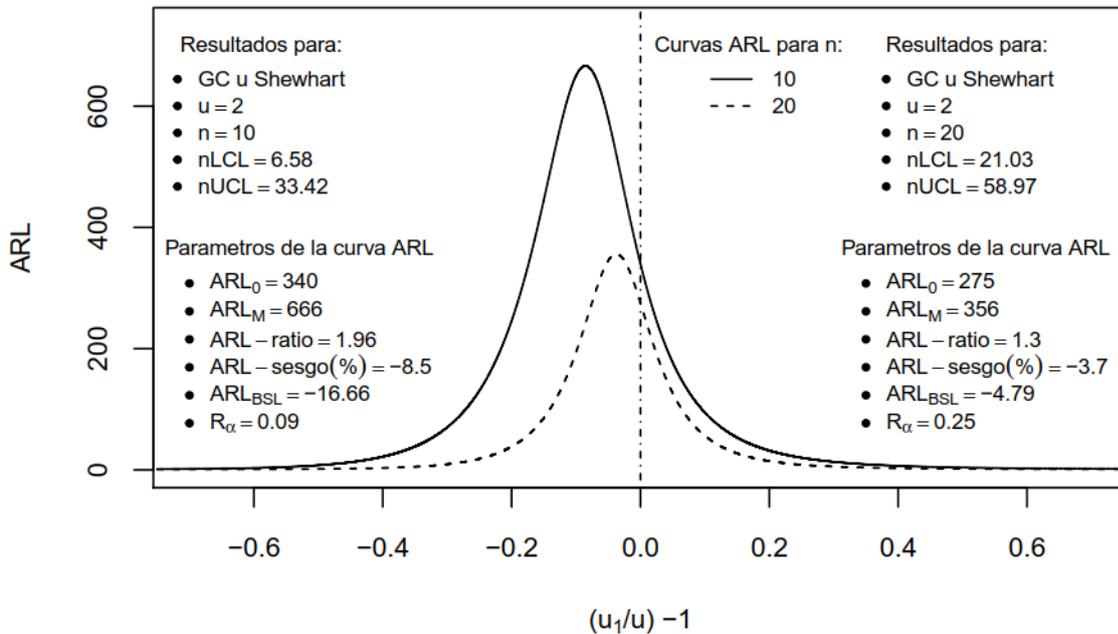


GC u Shewhart – Comportamiento del ARL_0 para $u = 2$

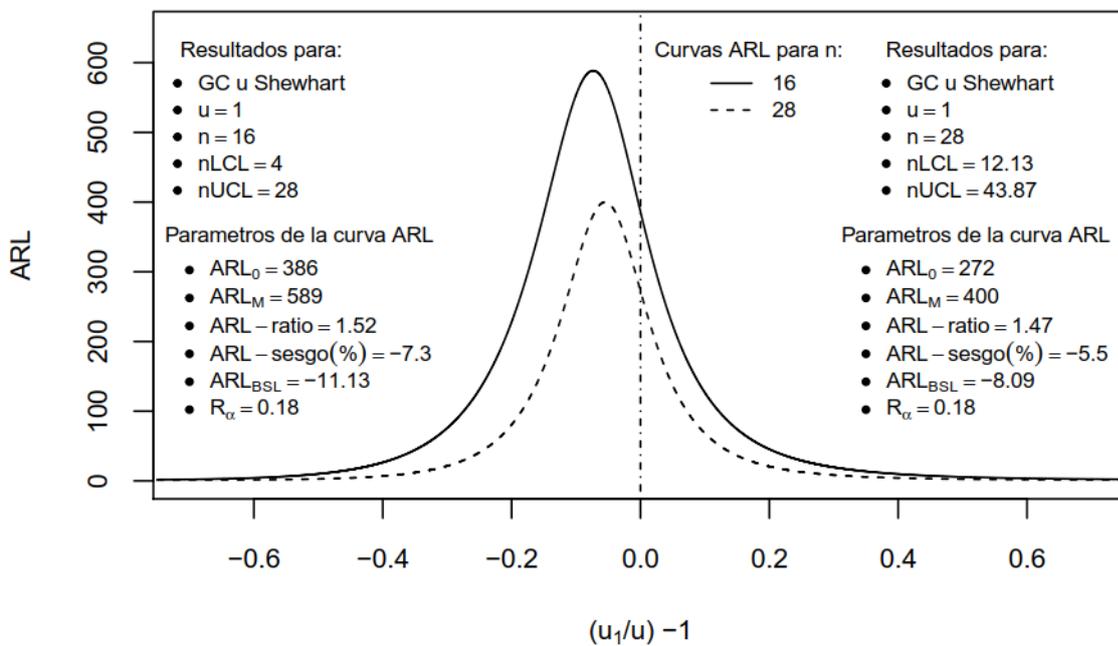
GC u Shewhart - Comportamiento real de la colas alfa para $u = 3$ GC u Shewhart – Comportamiento del ARL_0 para $u = 3$

8.8 Anexo 8

8.8.1 Gráfico de control *u* Shewhart: Ejemplos de Curvas ARL (ver sección 4.2.6)



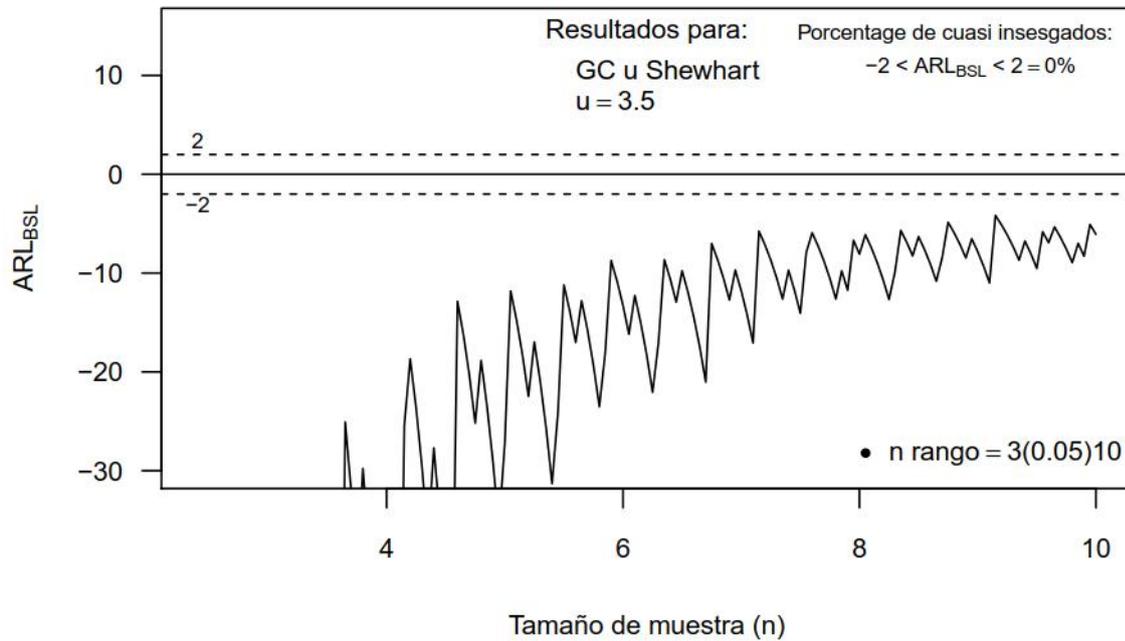
GC *u* Shewhart: Curvas ARL para $u=2$ con $n=10$ y 20



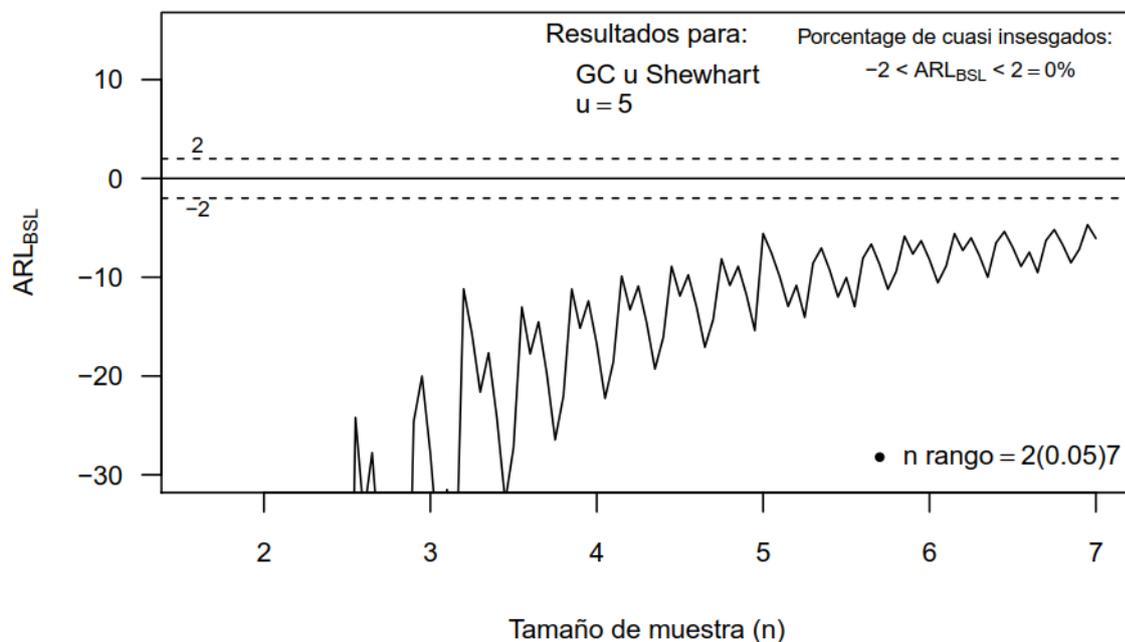
GC *u* Shewhart: Curvas ARL para $u=1$ con $n=16$ y 28

8.9 Anexo 9

8.9.1 Gráfico de control *u* Shewhart: Ejemplos de resultados ARL_{BSL} (ver sección 5.6.1)



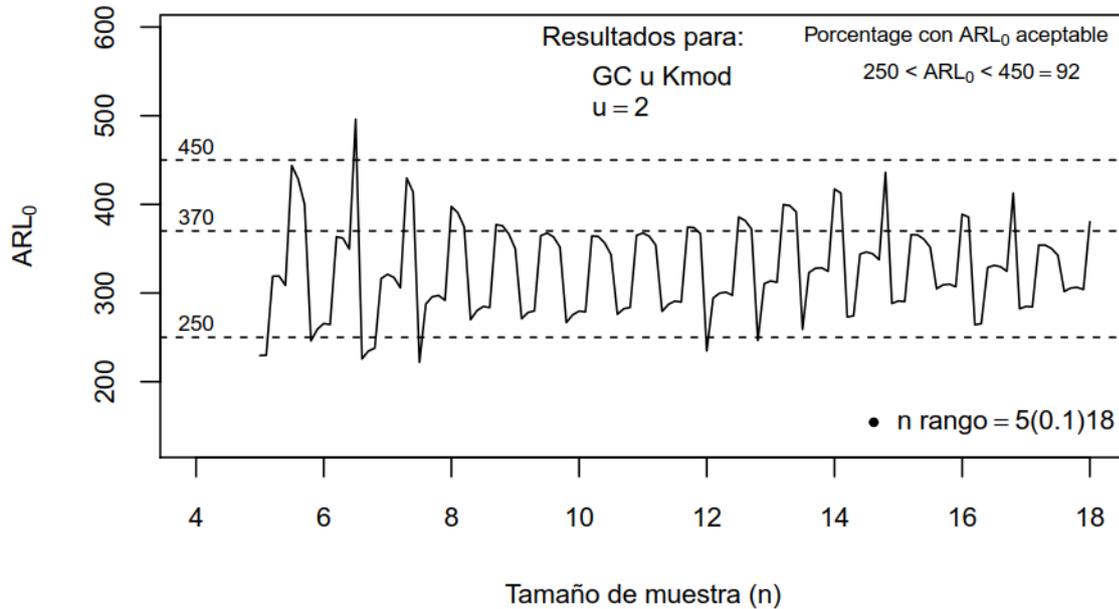
GC u Shewhart: Resultados del análisis ARL_{BSL} para $u = 3.5$



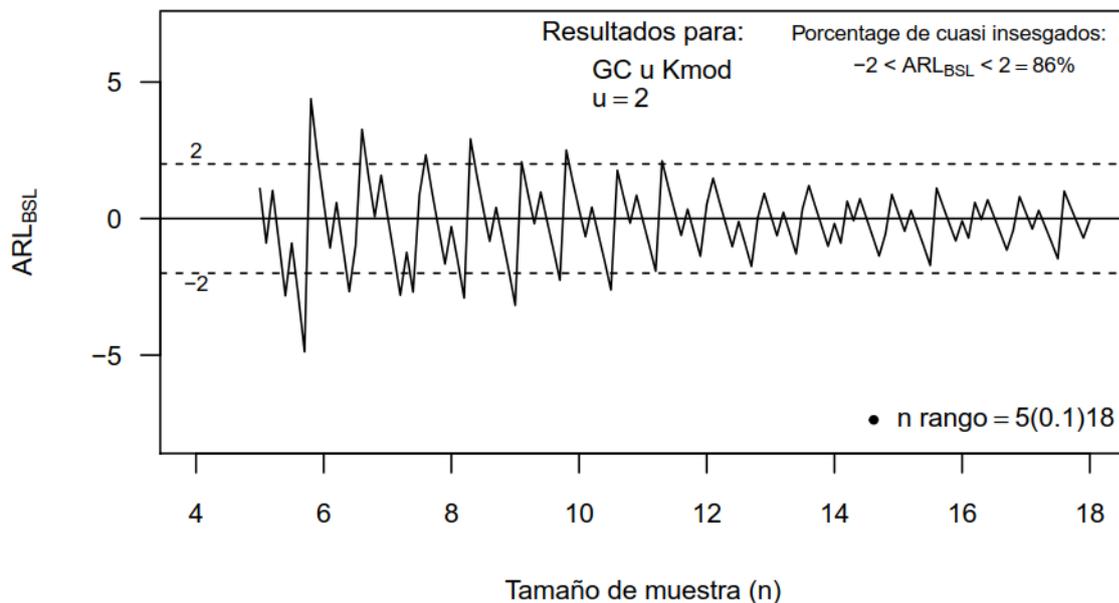
GC u Shewhart: Resultados del análisis ARL_{BSL} para $u = 5$

8.10 Anexo 10

8.10.1 Gráfico de control u K_{mod} : Ejemplos de resultados ARL_0 y ARL_{BSL} (ver sección 5.6.3.2)



GC u Kmod: Comportamiento del ARL_0 para $u = 2$



GC u Kmod: Comportamiento del ARL_{BSL} para $u = 2$

8.11 Anexo 11

8.11.1 Algunos códigos R utilizados en la investigación

```

# R codes for "Kmod p Chart"
#:#2018:.....

## Function for computing the chart's control limits
#.....

Kmod.CI<-function(p){

  n=n.optimo
  q=(1-p)

  kmod.lcl=3-((1.6)/((sqrt(n*p*q))))
  kmod.ucl=3+((1)/((sqrt(n*p*q))))

  ucl1<-p+kmod.ucl*sqrt((p*q)/n)
  lcl1<-p-kmod.lcl*sqrt((p*q)/n)

  CI_List<-list("ucl"=ucl1,"lcl"=lcl1)

  return(CI_List)
}

#.....

#To establish n range values
p=0.05

dmax=0.99*p

n1.min=round((9*p*(1-p))/(dmax^2),0)
dmin=p/2
n1.max=round((9*p*(1-p))/(dmin^2),0)
n1.max

#Sample size range
n.rangeStep=1
n2=seq(n1.min,n1.max,n.rangeStep)
MinSampleSize=n1.min # requerido para que salga en los graficos en n minim
o analizado

#.....

#Loop for computing the relevant ARL parameters

#::START LOOP:::

```

```

CI="GC p Kmod" # para que salga en los graficos
method= Kmod.CI
#.....
LowerTailAlfa<-0
UpperTailAlfa<-0
AlphaTailRatio<-0
ARL0<-0
ARLcurveMax<-0
ARLBiasPercent<-0

library(progress)# usado para ver el progreso del Loop
pb <- progress_bar$new(format = " Loop Progress [:bar] :percent in :elapse
d",total = length(n2), clear = FALSE, width= 60)

for(i in 1:length(n)){
# To check Loop progress

pb$tick()# para ver el progreso del Loop

n.optimo=n[i]
# To calculate Limits of Chart
CCWald.CI.all<- lapply(p,method)
lcl.CCWald<-sapply(CCWald.CI.all,`[[`, 2)
ucl.CCWald<-sapply(CCWald.CI.all,`[[`, 1)

n.optimoUCL<-(ucl.CCWald*n.optimo)
n.optimoLCL<-(lcl.CCWald*n.optimo)
# Calculate the alpha tails
UpperTailAlfa[i]<-1-pbinom(n.optimoUCL,size =n.optimo ,prob =p)
LowerTailAlfa[i]<-pbinom(n.optimoLCL,size =n.optimo ,prob =p)

# Para calcular el Alpha tails Ratio
AlphaTailRatio[i]<-LowerTailAlfa[i]/UpperTailAlfa[i]

# Para calcular el ARL0
ARL0[i]<-1/(UpperTailAlfa[i]+LowerTailAlfa[i])

#To compute the values for the ARL Curve

p.shift=seq(0.25*p,4*p,0.00001)
if(n.optimoLCL<=0){p.shift=seq(p,3*p,0.00001)}else{p.shift} # Si los limit
es son negativos no se dan los valores ARL
# Valores normalizados de p-shift
p.shift.normal.mid.p<-((p.shift/p)-1)# normalized percentage shift values
CBeta.CCWald<-pbinom(n.optimoUCL,size =n.optimo ,prob =p.shift)-pbinom(n.o
ptimoLCL,size =n.optimo ,prob =p.shift)
CFAR.CCWald<-1-CBeta.CCWald
CARL.CCWald<-1/CFAR.CCWald

CCWald_optimo.ARL.Results<-data.frame(p.shift,p.shift.normal.mid.p,CFAR.CC
Wald,CARL.CCWald)

#To compute the ARLmax y el ARL-sesgo(%)

```

```

CCWald.MaxARL<-max(CCWALD_optimo.ARL.Results$CARL.CCWald)
ARLcurveMax[i]<-CCWald.MaxARL # This is the ARLmax de la curva ARL

CCWald.ARL_Bias<-subset(CCWALD_optimo.ARL.Results, CARL.CCWald==CCWald.Max
ARL, select = c(p.shift.normal.mid.p))

Percent.CCWald.ARL.Bias<-round(CCWald.ARL_Bias[,1]*100,6) # This is the AR
L bias (%)
ARLBiasPercent[i]<-Percent.CCWald.ARL.Bias
}
#::: END LOOP ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
#
# To round the alpha values
AlphaTailRatio.r<-round(AlphaTailRatio,6)
# To round ARL0 values
ARL0.r<-round(ARL0,2)
# To round ARLmax values
ARLcurveMax.r<-round(ARLcurveMax,2)
# To obtain the ARL-ratio
ARL.ratio= ARLcurveMax/ARL0

QuasiUnbiasedResults.1<-data.frame(n2,AlphaTailRatio.r,ARL0.r,ARLcurveMax.
r,ARL.ratio,ARLBiasPercent)

data1<-data.frame(n2,UpperTailAlfa,LowerTailAlfa,ARL0.r)

#

```

```

# To plot the Alpha tails vs. sample size

plot(n2,UpperTailAlfa,type="l",cex=0.5,col="black",
      if (T){xlim = c(MinSampleSize*0.8,n1.max)},
      if(T){ylim =c(ifelse(min(UpperTailAlfa)<min(LowerTailAlfa),
                          min(UpperTailAlfa),min(LowerTailAlfa))*0.1,
                     ifelse(max(UpperTailAlfa)>max(LowerTailAlfa),
                          max(UpperTailAlfa),max(LowerTailAlfa))*1.2},
      xlab = "Tamaño de muestra (n)", ylab = "Probabilidad de error Tipo I"
      ,
      cex.axis=0.8, cex.lab=0.9)
abline(h=0.00135,lty=2)
text(MinSampleSize*0.8,0.00135, "0.00135",col = "black",adj = c(0.3,-0.4),
srt = 0,cex = 0.7)
lines(n2,LowerTailAlfa,lty=2,col="black",cex=0.5) legend(x="topright",lege
nd = c("Cola alfa superior","Cola alfa inferior"),
      lty=c(1,2),col=c("black","black"), bty= "n",cex=0.8)

legend(x="top",title="Resultados para:",legend = c(as.expression(bquote(.
(CI))),
                                                    as.expression(bquote(p=.(p
))))),bty = "n",cex=0.8)
legend(x=n1.max*0.75,y=mean(UpperTailAlfa)*0.7,inset=.01,title="",

```

```

        legend=c(as.expression(bquote("n rango"==.(n1.min)*("*. (n.rangeSte
p)*)"*.(n1.max))))),
        pch=c(20,20), bty = "n",cex=0.8)

# End plot of Alpha tails
#

```

```

# To plot the ARL0 vs. n
plot(n2,ARL0,type = "l",pch=16,cex=0.2,
     if (T){xlim = c(0.95*MinSampleSize,max(n2))},
     if(T){ylim = c(min(ARL0)*0.5,max(ARL0)*1.2)},
     xlab = "Tamaño de muestra (n)", ylab =expression(paste("ARL"[0])),
     cex.axis=0.8, cex.lab=0.9)
abline(h=370,lty=2)
text(0.9*MinSampleSize,370, "370",col = "black",adj = c(0.5,-0.4),srt = 0,
     cex = 0.7)

abline(h=450,lty=2)
text(0.9*MinSampleSize,450, "450",col = "black",adj = c(0.5,-0.4),srt = 0,
     cex = 0.7)

abline(h=250,lty=2)
text(0.9*MinSampleSize,250, "250",col = "black",adj = c(0.5,-0.4),srt = 0,
     cex = 0.7)

legend(x="top",title="Resultados para:",
       legend = c(as.expression(bquote(.(CI))),
                 as.expression(bquote(p==.(p))))),
       bty = "n",cex=0.8)
legend(x=n1.max*0.75,y=min(ARL0)*1.2,inset=.01,title="",
       legend=c(as.expression(bquote("n rango"==.(n1.min)*("*. (n.rangeSte
p)*)"*.(n1.max))))), pch=c(20,20), bty = "n",cex=0.8)

# End plot of ARL0
#

```

```

# ARL(BSL) vs Sample size

Severity.level<-(ARL.ratio*ARLBiasPercent)
ARL.bs1_data<-data.frame(n2,ARL.ratio,Severity.level,ARL0)

plot(n2,Severity.level,
     las=1,
     type="l",cex = 0.5,
     ylim = c(-15,10),
     xlim = c(min(n2)*0.95,max(n2)),
     log = "x",
     xlab="Tamaño de muestra (n)",
     ylab=expression(paste("", "ARL"[BSL])),

```

```
      cex.axis=0.8, cex.lab=0.9)
abline(h=0,lty=1)
abline(h=2,lty=2)
text(min(n2)*0.95,2,"2",col = "black",adj = c(0.5,-0.4),srt = 0,cex = 0.7)
abline(h=-2,lty=2)
text(min(n2)*0.95,-2,"-2",col = "black",adj = c(0.5,1.2),srt = 0,cex = 0.7
)

legend(x="top",title="Resultados para:",
      legend = c(as.expression(bquote(. (CI))),
                as.expression(bquote(p==.(p)))),
      bty = "n",cex=0.8)
legend(x="bottomright",inset=.01,title="",
      legend=c(as.expression(bquote("n rango"==.(n1.min)*("*. (n.rangeSte
p)*")"*.(n1.max)))),
      pch=c(20,20), bty = "n",cex=0.8)

#
```
