

Índice General

1	Estadística Aplicada al Control de Calidad	5
1.1	Estadística Descriptiva	6
1.1.1	Definiciones Básicas	6
1.1.2	Tipos de Datos	6
1.1.3	La Distribución de Frecuencias y el Histograma	7
1.1.4	Diagrama de Pareto	8
1.1.5	Diagrama de Causas y Efectos o Diagrama de Raspa de Pescado	11
1.1.6	Resumen Numérico de los Datos	12
1.2	Probabilidad	13
1.2.1	Distintas concepciones de Probabilidad. Definición axiomática	14
1.2.2	Variables aleatorias	16
1.3	Distribuciones de probabilidad	17
1.3.1	Distribuciones Discretas Importantes	19
1.3.2	Distribuciones Continuas Importantes	24
1.3.3	Teorema Central del Límite	29
1.3.4	Jerarquía de Aproximaciones	29
1.4	Inferencia Estadística	29
1.4.1	Estadísticos Muestrales y Distribuciones de Muestreo	31
1.4.2	Estimación de Parámetros	36
1.4.3	Contraste de hipótesis	45

1.5	Ejercicios	51
2	Control Estadístico de Procesos	57
2.1	Introducción Histórica	57
2.2	¿En qué consiste el Control de Calidad?	61
2.3	Algunas Definiciones en Control de Calidad	62
2.4	Causas Naturales y Atribuibles	63
2.5	Funcionamiento de un Gráfico de Control	64
2.5.1	Principios Básicos	64
2.5.2	Patrones de Comportamiento	68
2.5.3	Diseño de un Gráfico de Control	68
2.6	Límites de aviso	76
2.7	Curva Característica Operativa	78
2.7.1	La Curva OC y el Tamaño Muestral	81
2.7.2	La Curva OC y la Longitud Media de Racha	83
2.8	Ejercicios	86
3	Gráficos de Control de Atributos	91
3.1	Introducción	91
3.2	Gráfico de Control para la Fracción de Disconformes: Gráfico p	92
3.2.1	Tamaño Muestral Constante	96
3.2.2	Tamaño Muestral Variable	99
3.2.3	Gráfico p Normalizado	103
3.3	Gráfico np	104
3.4	Gráficos de Control de Disconformidades	109
3.4.1	Gráfico de Control del Número de Disconformidades en una Unidad de Inspección: Gráfico c	109
3.4.2	Gráfico c Normalizado	111
3.4.3	Gráfico de Control del Número de Disconformidades por Unidad: Gráfico u	111
3.5	Sistema de Calificación de Deméritos	117
3.6	Ejercicios	123

4	Gráficos de Control para Características Variables de Calidad	135
4.1	Introducción	135
4.2	Gráficos de Control	135
4.2.1	Parámetros Conocidos o Estándares	136
4.2.2	Parámetros Desconocidos	141
4.3	Interpretación de los Gráficos para la Media y la Variabilidad	149
4.4	Ejercicios	159
5	Procedimientos Especiales de Control de Procesos	165
5.1	Otros Gráficos de Control de Variables	165
5.1.1	Gráfico de Control para Medidas Individuales	165
5.1.2	Gráfico de Control de Media y Varianza	170
5.2	Gráficos con Memoria	173
5.2.1	Gráfico de la Media Móvil (<i>MA</i>)	174
5.2.2	Gráfico de la Media Móvil Geométrica (<i>EWMA</i>)	177
5.2.3	Gráficos de Sumas Acumuladas (<i>CUSUM</i>)	180
5.3	Ejercicios	187
6	Análisis de la Capacidad del Proceso	193
6.1	Introducción	193
6.2	Estimación de la Variabilidad del Proceso	194
6.3	Límites de Control, Límites de Tolerancia Natural y Límites de Especificación	195
6.4	Estimación de la Proporción de No Conformes	197
6.5	Índices de Capacidad	197
6.5.1	Índice de Funcionamiento	201
6.5.2	Índices de Capacidad Unilaterales	203
6.5.3	Índice de Capacidad Recíproco C_r	204
6.5.4	Índice de Capacidad Modificado C_{pm}	205
6.5.5	Coeficiente n -sigma	206
6.6	Errores de Medida e Índices de Capacidad	208
6.7	Capacidad del Proceso: Atributos	210

6.7.1	Proporción de Disconformes	211
6.7.2	Número de Disconformidades	211
6.8	Ejercicios	213
7	Introducción al Muestreo para la Aceptación	217
7.1	Introducción	217
7.2	Formación de los lotes	219
7.3	Planes de muestreo único para atributos	220
7.3.1	Introducción	220
7.3.2	Curva Característica Operativa	220
7.3.3	Probabilidad de aceptación de un lote	223
7.3.4	Determinación de un Plan de Muestreo	227
7.3.5	Sistemas de muestreo para la aceptación	228
7.4	Inspección de rectificación	232
7.5	Muestreo doble y múltiple	235
7.5.1	Introducción	235
7.5.2	Curvas OC para muestreo doble y múltiple	238
7.5.3	Tamaño Muestral Medio	241
7.6	El modelo MIL-STD-105E	243
7.6.1	Desarrollo histórico	243
7.6.2	MIL-STD-105E	243
7.6.3	Implementación	247
7.7	Planes de muestreo para variables	250
7.7.1	Introducción	250
7.7.2	Muestreo para la Aceptación con Variables	251
7.8	El Modelo <i>MIL-STD-414</i>	255
7.8.1	Introducción	255
7.8.2	<i>MIL-STD-414</i>	255
7.8.3	Implementación	257
7.9	Ejercicios	261
A	Tablas del <i>Mil-STD-505E</i>	263

Capítulo 1

Estadística Aplicada al Control de Calidad

Con el término *Estadística* se recogen una gran diversidad de técnicas encaminadas a analizar la información mediante la observación y la experimentación. Es la ciencia que estudia los procedimientos destinados a la recogida, resumen, análisis e interpretación de un conjunto de datos, así como los conducentes a la obtención de deducciones científicas a partir de ellos. Dentro de la Estadística cabe destacar tres ramas:

- *Estadística Descriptiva*, que se ocupa del tratamiento de sucesos que ya han ocurrido (datos).
- *Inferencia Estadística*, que generaliza a toda la población la información obtenida a partir de una muestra reducida.
- *Cálculo de Probabilidades*, que está a caballo entre las dos anteriores y se trata de un conjunto de proposiciones que permiten medir la aleatoriedad de un fenómeno a través de la asignación de un modelo matemático a dicho fenómeno.

En este capítulo se pretende recordar e introducir algunos conceptos básicos de *Estadística Descriptiva*, *Probabilidad* e *Inferencia Estadística* que serán de utilidad en el posterior desarrollo del Control Estadístico de Procesos.

1.1 Estadística Descriptiva

Como ya se ha comentado, la *Estadística Descriptiva* se encarga de recoger, clasificar y resumir los datos de un experimento en unas pocas medidas, de manera que en ellas esté contenida la máxima información sobre ellos.

A continuación, se estudian algunos conceptos propios de la Estadística Descriptiva.

1.1.1 Definiciones Básicas

Algunas definiciones importantes son:

- *Población*: Es el conjunto de elementos o individuos objeto de estudio con una característica común. Debe estar perfectamente determinada, de modo que pueda discernirse claramente si un individuo pertenece o no a la misma.
- *Individuo*: Es cada uno de los elementos de la población.
- *Muestra*: Es un subconjunto pequeño de la población, que debe ser representativa de toda ella. Se van a realizar estudios sobre muestras.
- *Carácter*: Es una propiedad común que se desea estudiar de los elementos de la población. Éste puede ser cuantitativo o cualitativo.
- *Modalidad*: Es cada uno de los posibles estados que puede presentar cada carácter, de modo que han de ser excluyentes y exhaustivos, es decir, cada individuo presenta una y solo una de las modalidades del carácter.
- *Variable*: Se utiliza esta denominación para referirse al carácter estudiado, cuando éste es cuantitativo.

1.1.2 Tipos de Datos

Se pueden distinguir dos tipos de datos:

1. *Datos de tipo cualitativo o atributos*, que son aquellos que se refieren a una cualidad de la población estudiada y que no pueden ser medidos

numéricamente. Como ejemplo, se puede considerar un artículo que se inspecciona y se clasifica como conforme o disconforme.

2. *Datos de tipo cuantitativo*, que son aquellos que pueden ser medidos numéricamente. A su vez, los datos de tipo cuantitativo se dividen en:

- (a) *Discretos*, que son aquellos que sólo toman valores aislados. Como ejemplo, se puede considerar el número de artículos defectuosos tras inspeccionar un determinado lote de artículos.
- (b) *Continuos*, que son aquellos que pueden tomar cualquier valor dentro de un intervalo, por ejemplo, la longitud de un determinado componente.

1.1.3 La Distribución de Frecuencias y el Histograma

Cuando se tiene un proceso de fabricación de un determinado producto, no hay dos unidades del producto que sean totalmente idénticas respecto de la variable objeto de estudio, puesto que es inevitable que haya alguna variación entre ellas. Por ejemplo, el peso neto de un cierto producto fabricado varía ligeramente en relación con otro. La Estadística analiza los datos obtenidos y saca conclusiones a partir de éstos, teniendo en cuenta las variaciones de los datos.

Una *distribución de frecuencias* consiste en un ordenamiento de los datos según su magnitud, de manera que cada modalidad que toma la variable estudiada viene representada por su frecuencia, es decir, el número de veces que se ha obtenido en la muestra estudiada el valor o conjunto de valores representado por la modalidad.

La distribución de frecuencias se representa gráficamente mediante un *histograma*, en el que la altura de sus rectángulos se corresponde con la frecuencia de las observaciones. El histograma es una representación visual de los datos en la que pueden observarse más fácilmente tres propiedades:

1. Forma de la distribución de frecuencias.
2. Medidas de acumulación o tendencia central de la variable estudiada.
3. Dispersión o variabilidad de la variable.

Hay muchos métodos para construir histogramas. Cuando los datos son numerosos, es muy útil utilizar intervalos de valores como modalidades de la variable estudiada. En general:

- Se recomienda utilizar entre 4 y 20 grupos (o intervalos); a menudo conviene elegir un número total de grupos aproximadamente igual a la raíz cuadrada del tamaño de la muestra.
- Los intervalos deben tener amplitud uniforme (en caso contrario, la altura de los rectángulos sería la densidad de frecuencias de cada intervalo, es decir, el cociente entre la frecuencia y la amplitud del intervalo, que mide la frecuencia por unidad de medida).
- Se empieza con un límite inferior del primer intervalo ligeramente menor que el valor más pequeño de los datos.

El agrupamiento de los datos condensa los datos originales, lo que da como resultado una pérdida de detalle. Así, cuando el número de observaciones es relativamente pequeño, o cuando las observaciones sólo toman pocos valores, puede construirse el histograma a partir de la distribución de frecuencias de los datos sin agrupar, utilizando modalidades simples.

En el Control de Calidad la interpretación de los histogramas puede ayudar a la comprensión de lo que está ocurriendo con el proceso.

Ejemplo 1 *Se considera la variable longitud de los tornillos fabricados. Se toma una muestra de 20 tornillos de 55 mm y se mide su verdadera longitud en mm., obteniendo los siguientes resultados, que vienen dados por la distribución de frecuencias de la Tabla 1.1.*

Longitud	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
Frecuencia	1	2	3	3	2	1	2	3	2	1

Tabla 1.1: Frecuencia de la longitud de los tornillos fabricados.

Su histograma de frecuencias viene dado por la Figura 1.1.

1.1.4 Diagrama de Pareto

Otro tipo de diagrama de barras diferente al histograma es el diagrama de Pareto, si bien su construcción es similar. En este caso también se

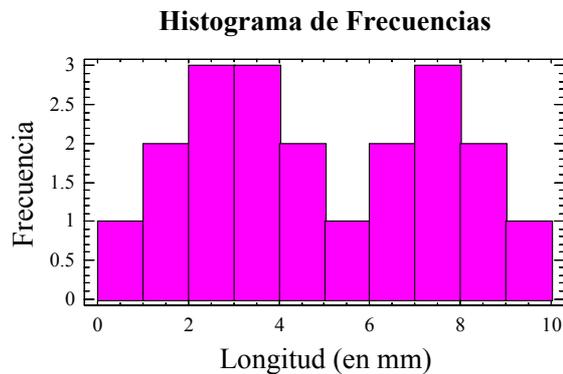


Figura 1.1: Histograma de la longitud de los tornillos fabricados

contabiliza la frecuencia con la que se da una clase, pero en el eje horizontal no aparece una escala continua graduada, sino que se representan varias categorías de interés. Los diagramas de Pareto fueron introducidos por Juran (1964) y se basan en datos de atributos, no de variables numéricas.

Para su construcción, se comienza determinando un área de interés para representar en el eje horizontal y se subdivide en distintas modalidades. Después se cuenta el número de veces que ocurre cada modalidad y las frecuencias resultantes se dibujan como las alturas de las correspondientes barras del diagrama. Estos diagramas centran la atención en factores importantes o problemas que afectan al proceso cuando las barras (y las modalidades asociadas) se reorganizan, por lo tanto la barra más alta se coloca en el lado izquierdo del diagrama. El diagrama de Pareto ayuda a distinguir los problemas importantes de los problemas poco importantes.

Se llaman diagramas de Pareto porque normalmente muestran los efectos de Pareto, que hacen referencia al hecho de que, en la mayoría de los casos, sólo un pequeño grupo de barras tiende a acumular la mayoría de los datos.

Ejemplo 2 *En un proceso de fabricación, los productos terminados pasan por una inspección final antes de ser vendidos. Durante esta inspección, se encuentran los productos defectuosos y se recogen los distintos tipos de defectos. Para una semana concreta se obtuvieron los que aparecen en la*

Tabla 1.2. La categoría “otros defectos” convencionalmente se coloca la última, ya que contiene muchos pequeños fallos.

Tipo de Defecto	Nº de productos rechazados
A	6
B	2
C	7
D	2
E (otros)	2

Tabla 1.2: Tipo de defecto de los artículos rechazados.

Diagrama de Pareto para el tipo de defecto

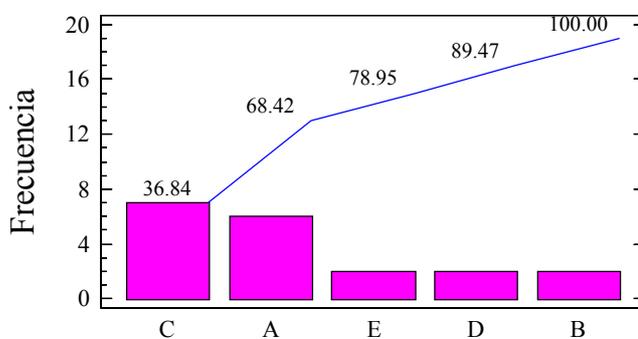


Figura 1.2: Diagrama de Pareto del tipo de defecto de los artículos rechazados

En la Figura 1.2 se aprecia que la mayoría de los defectos son del tipo C y A.

Como usualmente se desea convertir rápidamente las frecuencias de las modalidades en sus porcentajes relativos comparando con todos los datos, los diagramas de Pareto normalmente tienen dos escalas verticales. La de la izquierda, que representa las frecuencias, y la de la derecha, que representa el rango del porcentaje. Como una ayuda visual se representa una línea de

frecuencias acumuladas para juzgar fácilmente la contribución adicional de las sucesivas modalidades.

1.1.5 Diagrama de Causas y Efectos o Diagrama de Raspa de Pescado

El diagrama de causas y efectos es un método gráfico de flechas en el que el problema que se está estudiando se dibuja en la cabeza del diagrama. Después se sugieren las posibles causas de este problema y se dibujan líneas adicionales, una para cada una de estas causas. Para comenzar este tipo de diagramas es conveniente comenzar con unas pocas causas. Después, considerando cada causa como el efecto de otras causas, se rellena el diagrama de forma iterativa con más y más niveles. El aspecto de espina de pescado del diagrama le da el otro nombre por el que es conocido. Lo más importante a considerar cuando se hace este diagrama es que los elementos que se introducen en cada rama se consideran la causa de la rama a la que está dirigida. Cada rama, por lo tanto, es tanto una causa como un efecto, y el diagrama ilustra de forma clara sus interconexiones.

Los diagramas de causas y efectos se utilizan con más frecuencia cuando la fuente del problema no se conoce y es necesario reflexionar para descubrirla. Esto ocurre por ejemplo, cuando se utilizan diagramas de Pareto para resolver problemas. Los sucesivos niveles de estratificación requieren el estudio de las raíces del problema que indica el diagrama de Pareto, lo que no resulta fácil ya que no siempre es sencillo identificar las clases que hay que introducir en el diagrama de Pareto. En este caso los diagramas de Pareto y el de causas y efectos se utilizan subsecuentemente, esto es, el diagrama de Pareto inicial se emplea para identificar la primera modalidad a atacar, y después se utiliza un diagrama de causas y efectos para producir una lista de modalidades del siguiente diagrama de Pareto. De este último diagrama de Pareto, se elige la modalidad con más frecuencia y se selecciona para el estudio, y se vuelve a hacer un diagrama de causas y efectos para determinar las nuevas clases, y así sucesivamente. Esencialmente, la clase principal del diagrama de Pareto se utiliza como cabeza del subsecuente diagrama de causas y efectos, el cual identifica las modalidades para el siguiente diagrama de Pareto.

Ejemplo 3 *En la Figura 1.3 se tiene el diagrama de causas y efectos de las posibles razones por las que en el Ejemplo 2 se han obtenido los defectos*

del tipo C.

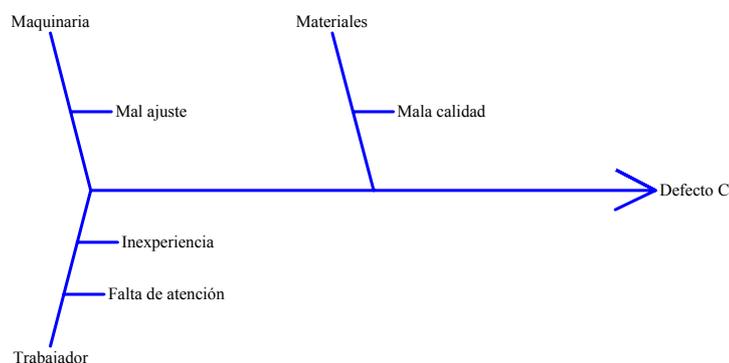


Figura 1.3: Diagrama de Causas y Efectos de los defectos de tipo C

1.1.6 Resumen Numérico de los Datos

Como se comentó en la Sección 1.1.3, el histograma es una representación visual de tres propiedades de datos muestrales: la forma de su distribución, su tendencia central y su dispersión o variabilidad. También es útil utilizar medidas numéricas para la tendencia central y la dispersión.

Sean x_1, x_2, \dots, x_n las observaciones de una muestra. Las medidas más destacadas son:

- Para la tendencia central, la *media muestral*, que se define como:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

La media muestral es el punto de equilibrio del histograma. Así, esta medida representa el centro de masa o centro de gravedad de los datos muestrales.

- La dispersión o variabilidad de los datos muestrales respecto de la media muestral se mide con la *varianza muestral*, que se define como:

$$S^2 = \frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Si no hay variabilidad en la muestra, entonces para cada observación muestral $x_i = \bar{x}$ y $S^2 = 0$. Generalmente, cuanto más grande sea la varianza muestral S^2 , tanto mayor será la dispersión de los datos y viceversa.

La unidad de la varianza muestral es el cuadrado de la unidad general de los datos. Esto es muchas veces inconveniente y difícil de interpretar, por lo que se suele preferir utilizar como medida de dispersión la raíz cuadrada positiva de S^2 , S , denominada *desviación típica muestral*, es decir:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

La ventaja principal de S es que se expresa en la unidad original de las mediciones. Un punto importante es que la desviación típica no refleja la magnitud de los datos muestrales, sino solamente su dispersión alrededor de la media.

Para calcular S es mejor utilizar (para cálculos manuales):

$$S = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)}.$$

- Otra medida útil para medir la variabilidad de los datos muestrales es el *recorrido o rango muestral*, que se define como la distancia entre el mayor y el menor valor de la variable, es decir:

$$R = x_{(n)} - x_{(1)}$$

donde $x_{(n)} = \max_{1 \leq i \leq n} x_i$ y $x_{(1)} = \min_{1 \leq i \leq n} x_i$.

1.2 Probabilidad

Se estudian ahora algunos aspectos importantes relativos a la Probabilidad. El *Cálculo de Probabilidades* se ocupa del estudio de los fenómenos aleatorios, es decir, de aquéllos cuya ocurrencia está sujeta al azar; y su objetivo es la modelización matemática de estos fenómenos.

1.2.1 Distintas concepciones de Probabilidad. Definición axiomática

Seguidamente se indican los distintos conceptos de Probabilidad que han existido a lo largo de la historia hasta llegar a la definición formal:

- *Concepción clásica:*

Se considera un experimento en el que hay un número finito de resultados n que son igualmente factibles y excluyentes (no se puede obtener más de un resultado de forma simultánea) y sea A un resultado posible del experimento. Entonces, la probabilidad del resultado A se puede definir como:

$$P[A] = \frac{\text{casos favorables}}{\text{casos posibles}} \equiv \text{Regla de Laplace.}$$

- *Concepción frecuentista:*

En muchas ocasiones los posibles resultados de un experimento no son igualmente factibles. Si se repite el experimento muchas veces bajo las mismas condiciones, la probabilidad de un resultado se aproxima por su frecuencia y, a medida que aumenta el número de veces que se realiza el experimento (N), esta frecuencia relativa se aproximará más a su verdadero valor. Por tanto se puede definir la probabilidad de un resultado A como:

$$P[A] = \lim_{N \rightarrow \infty} f_N(A)$$

donde $f_N(A) = \frac{n_i}{N}$.

- *Concepción subjetiva:*

Si no se está en uno de los casos anteriores, la probabilidad se puede interpretar como el grado de creencia o convicción respecto a la ocurrencia de un resultado. El problema importante que presenta es que dos personas pueden asignar distintas probabilidades a un mismo resultado.

Definición de Probabilidad:

- Se realiza un experimento aleatorio:
 1. Cada uno de los posibles resultados de dicho experimento recibe el nombre de suceso elemental.
 2. El conjunto formado por todos los sucesos elementales se llama espacio muestral del experimento aleatorio (Ω). Se distinguen dos tipos de espacios muestrales: discretos y continuos, que, a su vez, pueden ser finitos o infinitos. Por ejemplo, el conjunto {conforme, defectuoso} es un espacio muestral discreto y finito; el número de llegadas a un determinado servicio {1,2,3,...} es un espacio muestral discreto e infinito; la longitud de un tornillo que deba de estar entre 1.9 y 2.1 cm. es un espacio muestral continuo, puesto que está formado por todas las posibles longitudes entre esos dos valores. Todo espacio muestral continuo es infinito.
 3. Un suceso de un espacio muestral es cualquier subconjunto del espacio muestral.

Vemos a continuación la definición formal de Probabilidad:

Definición 4 *Dado Ω un espacio muestral y E un suceso cualquiera, se define la probabilidad sobre Ω como una aplicación $P : \Omega \rightarrow \mathbb{R}$ verificando:*

1. $P[E] \geq 0$
2. $P[\Omega] = 1$
3. Si E_1, E_2, \dots son infinitos sucesos disjuntos dos a dos ($E_i \cap E_j = \emptyset$, para todo $i \neq j$), entonces se verifica que la probabilidad de la unión de los sucesos es la suma de las probabilidades de los distintos sucesos:

$$P[E_1 \cup E_2 \cup E_3 \cup \dots] = P[E_1] + P[E_2] + P[E_3] + \dots$$

Se puede decir que la probabilidad de un suceso es un número real que mide la posibilidad colectiva de ocurrencia de los resultados de ese suceso. Las probabilidades de los distintos sucesos se van a asignar utilizando cualquiera de los métodos ya vistos con anterioridad.

Como consecuencia de la definición, se tienen las siguientes propiedades:

1. La probabilidad de cualquier suceso E está entre 0 y 1.
2. P es no decreciente.
3. Sea E un suceso cualquiera. Entonces, $P[\bar{E}] = 1 - P[E]$, donde \bar{E} es el suceso contrario o complementario de E .

Independencia de sucesos

Dados dos sucesos A y B , son independientes si y solo si la probabilidad de la intersección de dichos sucesos es igual al producto de sus probabilidades, es decir, $P[A \cap B] = P[A]P[B]$.

1.2.2 Variables aleatorias

Se pretende estudiar el comportamiento de fenómenos aleatorios. Para modelizar estos fenómenos y poderlos tratar adecuadamente se introduce el concepto de variable aleatoria. Para el estudio de los experimentos aleatorios se hace necesario introducir una función que haga corresponder un número a cada suceso elemental del fenómeno.

Se define el concepto de *variable aleatoria* de la siguiente forma:

Definición 5 Sea Ω un espacio muestral sobre el que está definida una función de probabilidad P , y sea X una función de valor real definida sobre Ω tal que transforme los resultados del espacio muestral (sucesos) en valores reales. Entonces $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria. Además, cada variable aleatoria va a tener asociada una función de probabilidad.

Ejemplo 6 Sea X la variable aleatoria que estudia el número de cruces obtenidas en el lanzamiento de una moneda tres veces. El espacio muestral está formado por los 8 posibles resultados del experimento, cada uno con una probabilidad de $(1/2)^3 = 1/8$.

$$\Omega = \{ccc, +cc, c+c, cc+, ++c, +c+, c++, +++\}.$$

Los valores que toma la variable aleatoria X son: $X : \Omega \rightarrow \mathbb{R}$

$$\begin{array}{cccc} ccc & \longrightarrow & 0 & \quad +cc \longrightarrow 1 \quad c+c \longrightarrow 1 \quad cc+ \longrightarrow 1 \\ ++c & \longrightarrow & 2 & \quad +c+ \longrightarrow 2 \quad c++ \longrightarrow 2 \quad +++ \longrightarrow 3. \end{array}$$

Entonces, la probabilidad de los distintos valores que toma la variable aleatoria es:

$$P[X = x] = \begin{cases} P[X = 0] = P[(ccc)] = 1/8 \\ P[X = 1] = P[(+cc) \cup (c+c) \cup (cc+)] = 3/8 \\ P[X = 2] = P[(++c) \cup (+c+) \cup (c++)] = 3/8 \\ P[X = 3] = P[(+++)] = 1/8 \\ P[X = x] = 0; \quad x \neq 0, 1, 2, 3. \end{cases}$$

Se definen dos tipos de variables aleatorias:

Definición 7 Una variable aleatoria es discreta si sólo puede tomar un conjunto numerable de valores aislados.

Definición 8 Una variable aleatoria es continua si puede tomar cualquier valor dentro de un intervalo.

La función de distribución de una variable aleatoria calcula probabilidades de que la variable aleatoria tome valores en determinados subconjuntos:

Definición 9 La función de distribución $F(x)$ de una variable aleatoria X se define como:

$$F(x) = P[X \leq x], \quad x \in \mathbb{R}.$$

Ejemplo 10 La función de distribución del Ejemplo 6 viene dada por:

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3. \end{cases}$$

1.3 Distribuciones de probabilidad

Una *distribución de probabilidad* es un modelo matemático que relaciona cada valor de la variable con la probabilidad de ocurrencia de este

valor en la población. Por ejemplo, se puede considerar la longitud de un tornillo como una variable aleatoria, porque toma valores diferentes en la población conforme a algún mecanismo fortuito, y entonces la distribución de probabilidad de dicha longitud describe la probabilidad de ocurrencia de cualquier valor en la población. Hay dos tipos de distribuciones de probabilidad:

1. *Distribuciones continuas*, en las que la variable estudiada se mide en una escala continua. En este caso, la distribución de probabilidad se define mediante una función de densidad $f(x)$, tal que:

- $f(x) \geq 0$, para todo x valor de la variable aleatoria
- $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Además, la probabilidad de que la variable aleatoria tome valores dentro de un intervalo $[a, b]$ viene dada por la expresión:

$$P[a \leq X \leq b] = \int_a^b f(x)dx.$$

Así, por ejemplo, la distribución de probabilidad de la longitud de los tornillos estudiados es continua.

2. *Distribuciones discretas*, en las que la variable solamente puede tomar ciertos valores concretos x_1, x_2, \dots , (como los números enteros $1, 2, \dots$). En este caso, la distribución de probabilidad viene definida por un conjunto de valores $P[X = x_i]$, tales que:

- $P[X = x_i] \geq 0$, para todo x_i valor de la variable aleatoria.
- $\sum_{i=1}^{\infty} P[X = x_i] = 1$.

Por ejemplo, la distribución del número de disconformidades o defectos en un determinado producto es una distribución discreta.

Definición 11 *La media poblacional (μ) de una variable aleatoria X se llama esperanza y mide el valor promedio de dicha variable. Se define como:*

$$E[X] = \begin{cases} \sum_{i=1}^{\infty} x_i P[X = x_i] & \text{para una variable aleatoria discreta} \\ E[X] = \int_{-\infty}^{+\infty} x \cdot f(x) dx & \text{para una variable aleatoria continua.} \end{cases}$$

Definición 12 La dispersión o variabilidad de una variable aleatoria X respecto de su media se expresa mediante la varianza poblacional (σ^2), definida como:

$$\sigma^2 = \text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

1.3.1 Distribuciones Discretas Importantes

En esta sección se examinan las siguientes distribuciones discretas que aparecen frecuentemente en el Control Estadístico de Calidad: *Hipergeométrica*, *Binomial*, de *Poisson* y *Binomial Negativa*

Distribución Hipergeométrica

Se considera un experimento en las siguientes condiciones:

- Se tiene una población finita con N artículos.
- k (donde $k \leq N$) de estos artículos están en una clase de interés.
- Se selecciona una muestra aleatoria de n artículos de la población sin reposición.
- Sea X la variable que estudia el número de artículos en la muestra que están en la clase de interés.

Entonces se dice que la variable aleatoria X sigue una distribución *Hipergeométrica* $H(N, n, k)$, cuya función de probabilidad está dada por:

$$P[X = x] = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \quad x = 0, 1, 2, \dots, \min(n, k).$$

Además, su esperanza y varianza vienen dadas por:

$$\begin{aligned} \mu &= \frac{nk}{N} \\ \sigma^2 &= \frac{nk(N-n)(N-k)}{N^2(N-1)}. \end{aligned}$$

La distribución Hipergeométrica es el modelo probabilístico apropiado para seleccionar una muestra aleatoria de n artículos, sin reposición, de un lote de N de ellos, de los cuales k son disconformes o defectuosos. En estas aplicaciones X suele representar el número de artículos disconformes encontrados en la muestra.

Ejemplo 13 *Se considera un lote que contiene 10 artículos, de los que 2 son disconformes. Si se seleccionan 3 artículos al azar, sin reposición, entonces la probabilidad de encontrar a lo sumo un artículo disconforme en la muestra viene dada por:*

X : “número de artículos disconformes encontrados en la muestra”, $X \sim H(10, 3, 2)$

$$P[X \leq 1] = P[X = 0] + P[X = 1] = \frac{\binom{2}{0} \binom{8}{2}}{\binom{10}{3}} + \frac{\binom{2}{1} \binom{8}{1}}{\binom{10}{3}} = 0.9333333.$$

Distribución Binomial

Se considera el siguiente experimento:

- Se realiza una sucesión de n pruebas independientes bajo las mismas condiciones.
- El resultado de cada prueba es un *éxito* o un *fracaso*.
- La probabilidad de éxito, p , es constante, cada vez que se realiza el experimento, puesto que se realiza de manera independiente.
- Sea X la variable aleatoria que estudia el número de éxitos obtenidos en las n pruebas.

Entonces, se dice que la variable aleatoria X se distribuye según una *Binomial* de parámetros n y p , $B(n, p)$ y su función de probabilidad viene dada por la expresión:

$$P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x} \quad x = 0, 1, 2, \dots, n.$$

Además, su esperanza y varianza son:

$$\begin{aligned} \mu &= np \\ \sigma^2 &= np(1 - p). \end{aligned}$$

La distribución Binomial se utiliza con frecuencia en el Control Estadístico de la Calidad, puesto que es el modelo probabilístico apropiado para muestrear una población infinitamente grande, donde p representa la fracción o proporción de artículos defectuosos en la población. En estas aplicaciones, X suele representar el número de artículos disconformes encontrados en una muestra aleatoria de tamaño n .

Ejemplo 14 Si la proporción de artículos disconformes es del 10%, ¿cual es la probabilidad de que en una muestra de tamaño 5 aparezcan dos artículos disconformes?

X : “número de artículos disconformes encontrados en la muestra”, $X \rightsquigarrow B(5, 0.10)$

$$P[X = 2] = \binom{5}{2} 0.1^2 (0.9)^3 = 0.0729.$$

Una variable aleatoria que aparece con frecuencia en el Control Estadístico de la Calidad es \hat{p} , el cociente del número de artículos defectuosos en la muestra entre el tamaño muestral, denominada *fracción muestral de defectuosos* o *fracción muestral de disconformes*:

$$\hat{p} = \frac{X}{n} \quad \text{con } X \rightsquigarrow B(n, p).$$

La distribución de probabilidad de \hat{p} se obtiene a partir de la Binomial, ya que

$$P[\hat{p} \leq a] = P\left[\frac{X}{n} \leq a\right] = P[X \leq na] = \sum_{x=0}^{[na]} \binom{n}{x} p^x (1-p)^{n-x}$$

donde $[na]$ denota al mayor entero menor o igual que na . Además, su esperanza y varianza son:

$$\begin{aligned} E[\hat{p}] &= E\left[\frac{X}{n}\right] = \frac{1}{n} E[X] = \frac{np}{n} = p \\ \text{Var}[\hat{p}] &= \text{Var}\left[\frac{X}{n}\right] = \frac{1}{n^2} \text{Var}[X] = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}. \end{aligned}$$

Distribución de Poisson

Sea una variable aleatoria X que representa el número de sucesos aleatorios independientes que ocurren a velocidad constante en el tiempo o el espacio, y sea λ el número medio de ocurrencias por unidad de tiempo o espacio.

Entonces, se dice que la variable aleatoria X sigue una distribución de *Poisson*, $P(\lambda)$, con función de probabilidad dada por:

$$P[X = x] = \frac{e^{-\lambda} \lambda^x}{x!} \quad x = 0, 1, 2, \dots$$

Además, su esperanza y varianza son iguales y coinciden con el valor del parámetro, esto es:

$$\begin{aligned} \mu &= \lambda \\ \sigma^2 &= \lambda. \end{aligned}$$

Una aplicación clásica de la distribución de Poisson en el Control Estadístico de Calidad es como un modelo del número de defectos o disconformidades que ocurren en una unidad del producto. En realidad, cualquier fenómeno aleatorio que ocurre por unidad (de área, de volumen, de tiempo...) muchas veces se puede aproximar por una distribución de Poisson.

Ejemplo 15 *El número medio de defectos que tiene un determinado producto fabricado es 6. Si tiene más de 8, el producto no podrá ser vendido. Entonces, la probabilidad de no vender un producto es:*

X : “número defectos en un artículo”, $X \sim P(6)$

$$P\{X > 8\} = \sum_{i=9}^{\infty} \frac{e^{-6} 6^i}{i!} = 0.1528.$$

Distribución Binomial Negativa

Sea un experimento en las siguientes condiciones:

- El resultado del experimento es *éxito* o *fracaso*.
- Se realiza una sucesión de pruebas independientes, cada una de ellas con probabilidad de éxito p constante, hasta obtener el éxito número k .

- Sea X la variable que mide el número de fracasos antes del éxito número k .

Entonces se dice que la variable aleatoria X tiene distribución de probabilidad *Binomial Negativa*, $BN(k, p)$, con función de probabilidad dada por:

$$P[X = x] = \binom{x + k - 1}{x} p^k (1 - p)^x \quad x = 0, 1, 2, \dots$$

Además, su esperanza y varianza vienen dadas por:

$$\begin{aligned} \mu &= \frac{k(1-p)}{p} \\ \sigma^2 &= \frac{k(1-p)}{p^2}. \end{aligned}$$

Esta distribución es útil como modelo estadístico fundamental para varios tipos de datos de conteo, como la ocurrencia de disconformidades en una unidad de un producto.

Se puede utilizar la variable aleatoria $Y = k + X$, definida como el número de realizaciones del experimento hasta obtener el éxito número k . La distribución de probabilidad de la variable aleatoria Y es conocida como distribución de Pascal.

Ejemplo 16 *La probabilidad de tener que revisar 12 productos para obtener el tercero disconforme, si la probabilidad de que un producto sea disconforme es $1/6$, viene dada por:*

Y: “número de artículos extraídos de la muestra antes de encontrar el tercero disconforme”.

X: “número de artículos conformes encontrados en la muestra antes del tercero disconforme”, $X \rightsquigarrow BN(3, 1/6)$

$$P[Y = 12] = P[X = 9] = \binom{9 + 3 - 1}{9} (1/6)^3 (5/6)^9 = 0.0493.$$

Existe una importante diferencia entre las distribuciones Binomial y Binomial Negativa, ya que la primera fija el tamaño muestral y observa el número de éxitos; mientras que la segunda fija el número de éxitos y observa el tamaño muestral.

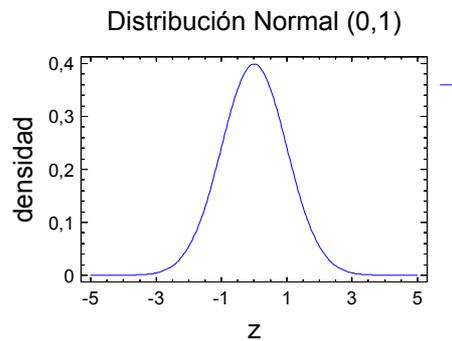


Figura 1.4: Función de densidad de la $N(0, 1)$

1.3.2 Distribuciones Continuas Importantes

Se estudian ahora las distribuciones continuas de probabilidad más útiles en Control de Calidad:

Distribución Normal

Si X es una variable aleatoria *Normal*, $N(\mu, \sigma)$, entonces su función de densidad viene dada por la expresión:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad -\infty < x < \infty, \mu \in \mathbb{R}, \sigma > 0$$

donde μ y σ^2 son la media y la varianza poblacionales, respectivamente.

Gráficamente, como se puede ver en la Figura 1.4, tiene forma de campana, de ahí que también se denomina *campana de Gauss*.

A partir de su gráfica, se puede observar que viene caracterizada por las siguientes propiedades:

- La probabilidad de ocurrencia aumenta a medida que la variable se acerca a su valor medio, mientras que disminuye a medida que se aleja de dicho valor medio.
- Las ocurrencias por encima y por debajo de dicho valor medio son igualmente probables, es decir, es simétrica respecto de su valor medio.

- La mayoría de las desviaciones de la media son pequeñas y es posible, aunque muy extraño, que se produzcan grandes desviaciones de la media.

Una propiedad de gran interés y utilidad es que si una variable aleatoria X sigue una distribución Normal de parámetros μ y σ , entonces la variable aleatoria X tipificada, es decir,

$$Z = \frac{X - \mu}{\sigma}$$

sigue una distribución Normal de parámetros 0 y 1.

En Control Estadístico de la Calidad, se utiliza la distribución Normal de tres formas:

- Cuando se miden ciertas características de calidad de un producto.
- Como una útil aproximación a la distribución Binomial, Poisson e Hipergeométrica.
- A partir del Teorema Central del Límite, que se estudiará posteriormente.

Ejemplo 17 *Una máquina llena una botella con una cantidad media de zumo. El volumen de zumo en cada botella tiene una distribución Normal con media 510 ml. y una desviación típica de 4 ml. ¿Cuál es la probabilidad de que una botella elegida al azar contenga menos de 500 ml de zumo?*

X : “contenido en ml. de la botella” sigue una distribución $N(510,4)$

Entonces:

$$P[X < 500] = P\left[Z < \frac{500 - 510}{4}\right] = P[Z < -2.5] = 0.0048.$$

Distribución Exponencial

La función de densidad de una variable *exponencial*, $Exp(\lambda)$, es:

$$f(x) = \lambda e^{-\lambda x} \quad x \geq 0$$

donde el parámetro $\lambda \geq 0$ es constante.

Además, su esperanza y varianza vienen dadas por:

$$\begin{aligned}\mu &= \frac{1}{\lambda} \\ \sigma^2 &= \frac{1}{\lambda^2}.\end{aligned}$$

La distribución exponencial se utiliza ampliamente en el campo de la ingeniería de fiabilidad como un modelo para el tiempo hasta el fallo de un componente o sistema. En estas aplicaciones, el parámetro se llama índice de falla del sistema y la media de la distribución se denomina tiempo hasta la falla.

Si el número de ocurrencias de un suceso sigue una distribución de Poisson con parámetro λ , entonces la distribución del intervalo entre dos ocurrencias consecutivas será una exponencial de parámetro λ .

Distribución Gamma

La función de densidad de una variable aleatoria *Gamma*, $G(\lambda, r)$, viene dada por:

$$f(x) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x} \quad x \geq 0$$

donde al parámetro $\lambda > 0$ se le llama parámetro de escala, al parámetro $r > 0$ parámetro de forma y Γ representa la función gamma.

Además, su esperanza y varianza vienen dadas por:

$$\begin{aligned}\mu &= \frac{r}{\lambda} \\ \sigma^2 &= \frac{r}{\lambda^2}.\end{aligned}$$

Si el parámetro r es un número entero, entonces la distribución Gamma es la suma de r distribuciones Exponenciales, distribuidas independiente e idénticamente, cada una con parámetro λ ; mientras que si $r = 1$, se tiene una distribución Exponencial de parámetro λ .

Ejemplo 18 *Se considera un sistema en el que hay dos fuentes de energía, de manera que si la primera funciona, la segunda se encuentra inactiva y*

cuando la primera falla, se activa automáticamente la segunda. Si cada componente tiene una duración descrita por una distribución exponencial con $\lambda = 10^{-4}$ horas, entonces la duración del sistema sigue una distribución $G(10^{-4}, 2)$. En consecuencia, el tiempo medio hasta el fallo es de $2 \cdot 10^4$ horas.

Distribución de Weibull

La función de densidad de una variable aleatoria de Weibull, $W(\gamma, \delta, \beta)$, es:

$$f(x) = \frac{\beta}{\delta} \left(\frac{x - \gamma}{\delta} \right)^{\beta-1} \exp \left[- \left(\frac{x - \gamma}{\delta} \right)^\beta \right] \quad x \geq \gamma$$

donde γ ($-\infty < \gamma < \infty$) es el parámetro de posición, $\delta > 0$ el parámetro de escala, y $\beta > 0$ el parámetro de forma.

Además, su esperanza y varianza vienen dadas por:

$$\begin{aligned} \mu &= \gamma + \delta \Gamma \left(1 + \frac{1}{\beta} \right) \\ \sigma^2 &= \delta^2 \left[\Gamma \left(1 + \frac{2}{\beta} \right) - \Gamma \left(1 + \frac{1}{\beta} \right)^2 \right] \end{aligned}$$

donde Γ representa la función gamma.

Esta distribución será útil cuando se mida la calidad de funcionamiento en términos de la fiabilidad del producto.

Distribuciones Asociadas a la Normal

Las siguientes distribuciones proceden de la distribución Normal:

Distribución χ^2 : Una distribución importante en el muestreo, definida en términos de la distribución Normal, es la distribución χ^2 . Si X_1, X_2, \dots, X_n son variables aleatorias Normales e independientemente distribuidas, $X_i \rightsquigarrow N(0, 1) \forall i = 1, \dots, n$, entonces la variable aleatoria:

$$\chi_n^2 = X_1^2 + X_2^2 + \dots + X_n^2 \rightsquigarrow \chi^2(n)$$

tiene distribución χ^2 con n grados de libertad, que es realmente una distribución $G(1/2, n/2)$.

La función de densidad de la distribución $\chi^2(n)$ se expresa como:

$$f(x) = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} (x)^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad x > 0.$$

La distribución tiene media $\mu = n$ y varianza $\sigma^2 = 2n$.

Distribución t -Student: Otra distribución útil en el muestreo es la distribución t -Student. Si $X \rightsquigarrow N(0, 1)$ y $Y \rightsquigarrow \chi^2(n)$ respectivamente y son independientes, entonces la variable aleatoria

$$T = \frac{X}{\sqrt{\frac{Y}{n}}} \rightsquigarrow t(n)$$

sigue una distribución t -Student con n grados de libertad.

La función de densidad de la distribución t -Student con n grados de libertad es:

$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2})} \left(\frac{t^2}{n} + 1\right)^{-\frac{(n+1)}{2}} \quad -\infty < t < +\infty$$

cuya media y varianza son $\mu = 0$ y $\sigma^2 = \frac{n}{n-2}$ para $n > 2$, respectivamente.

Si $n \rightarrow \infty$, la distribución t_n se aproxima a la $N(0, 1)$.

Distribución F -Snedecor: La última distribución muestral basada en el proceso Normal considerada es la distribución F -Snedecor. Si X e Y son dos variables aleatorias independientes tales que $X \rightsquigarrow \chi^2(n)$, $Y \rightsquigarrow \chi^2(m)$, entonces la razón:

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}} \rightsquigarrow F(n, m)$$

La función de densidad de la distribución F -Snedecor con n y m grados de libertad es:

$$g(f) = \frac{\Gamma(\frac{n+m}{2}) \left(\frac{n}{m}\right)^{\frac{n}{2}}}{\Gamma(\frac{n}{2}) \Gamma(\frac{m}{2})} \frac{f^{\frac{n}{2}} - 1}{\left[\left(\frac{n}{m}\right) f + 1\right]^{\frac{n+m}{2}}} \quad 0 < f < \infty.$$

1.3.3 Teorema Central del Límite

Se estudia ahora el *Teorema Central del Límite*, que es clave en el desarrollo de la Probabilidad.

Teorema 19 Sean X_1, \dots, X_n variables aleatorias independientes con media μ_i y varianza σ_i^2 , y sea $S = \sum_{i=1}^n X_i$, entonces:

$$\frac{S - E[S]}{\sqrt{\text{Var}(S)}} = \frac{S - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \xrightarrow{n \rightarrow \infty} N(0, 1)$$

Nota 20 Está claro que la aproximación mejora al aumentar n . En general, si la distribución de cada variable tiene forma aproximada de campana y es simétrica, el Teorema Central del Límite da una aproximación razonablemente buena para $n \geq 4$ y estas condiciones se cumplen con frecuencia en los problemas de Control de Calidad.

Corolario 21 Como caso particular, si todas las variables aleatorias son independientes e idénticamente distribuidas, con media μ y varianza σ^2 , entonces, el Teorema Central del Límite dice:

$$\frac{S - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{n \rightarrow \infty} N(0, 1)$$

1.3.4 Jerarquía de Aproximaciones

En el caso en el que no se pueda utilizar una determinada distribución de probabilidad debido a los altos valores de sus parámetros, se pueden realizar aproximaciones a otras distribuciones de probabilidad. Las más importantes se pueden ver en la Figura 1.5.

1.4 Inferencia Estadística

En el apartado anterior se analiza la utilización de las distribuciones de probabilidad para modelizar o describir una variable aleatoria, que representa la salida del proceso. En todos los casos se suponían conocidos los

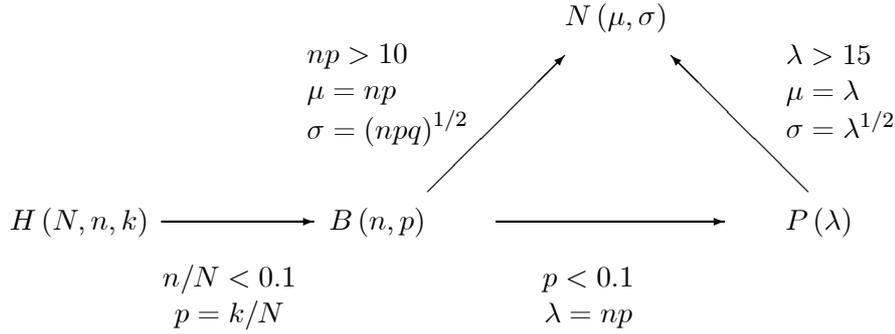


Figura 1.5: Aproximaciones entre distintas distribuciones

parámetros de la distribución de probabilidad y, por lo tanto, los parámetros del proceso. Esto no suele ocurrir en la realidad, ya que los parámetros de las distribuciones y, por tanto, del proceso suelen ser desconocidos. Por ejemplo, considérese la distribución Normal para modelizar el peso de un determinado producto mediante muestreo de un proceso de fabricación, con μ el primer parámetro de la distribución Normal. La interpretación física de μ es que se trata del verdadero valor medio de la variable aleatoria estudiada, es decir el verdadero valor medio de los pesos de los productos obtenidos en el proceso. Además, si se conociera el verdadero valor de μ y éste fuera relativamente constante en el tiempo, entonces se podría argumentar que los procedimientos formales de Control de Calidad son innecesarios.

En general, se desconocen los parámetros de la distribución de probabilidad de la variable aleatoria que controla el proceso y, además, éstos pueden cambiar a lo largo del tiempo. Por lo tanto, hay que desarrollar procedimientos para evaluar los parámetros de las distribuciones de probabilidad y resolver los problemas inferenciales u orientados hacia decisiones, relacionados con aquellos. Para ello, son útiles las técnicas estadísticas de estimación de parámetros y contrastes de hipótesis, por lo que se estudian algunos resultados elementales de la *Inferencia Estadística* y se indica su utilidad en los problemas de Control de la Calidad.

1.4.1 Estadísticos Muestrales y Distribuciones de Muestreo

El objetivo de la *Inferencia Estadística* es deducir o tomar decisiones con respecto a una población, a partir de los resultados obtenidos en una muestra seleccionada de dicha población.

Con frecuencia se supone que se emplean muestras aleatorias en el análisis. El calificativo *aleatorio* se aplica con frecuencia a cualquier método de selección no sistemático de una muestra. Si se seleccionan n variables aleatorias X_1, X_2, \dots, X_n , de manera que se distribuyen de forma independiente¹ e idéntica, se tiene una muestra aleatoria de tamaño n . Esta definición es adecuada para muestras aleatorias obtenidas de poblaciones finitas o infinitas con reposición. En un muestreo sin reposición de una población finita de N elementos, se dice que una muestra de n elementos es una muestra aleatoria si cada una de las $\binom{N}{n}$ muestras posibles tiene igual probabilidad de ser escogida.

Aunque la mayoría de los métodos que se estudian suponen el uso de un muestreo aleatorio, existen otras estrategias de muestreo que son útiles en el Control de Calidad. Hay que tener cuidado de utilizar un método de análisis que sea congruente con el diseño del muestreo. Las técnicas inferenciales proyectadas para muestras aleatorias pueden llevar a errores grandes cuando se aplican a datos obtenidos mediante otras técnicas de muestreo.

En la Inferencia Estadística no se trabaja directamente a partir de la muestra, sino que se utilizan cantidades calculadas a partir de las observaciones de la muestra denominadas *estadísticos muestrales*. Un estadístico muestral se define como una función de los datos muestrales que no contiene a los parámetros desconocidos. Por ejemplo, sea X_1, X_2, \dots, X_n una muestra aleatoria simple. Entonces las siguientes funciones son estadísticos muestrales:

1. La *media muestral*:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

¹El concepto de independencia de variables aleatorias se deriva del concepto de independencia de sucesos, por tanto, dos variables aleatorias son independientes si

$$P[(X \in A) \cap (Y \in B)] = P[X \in A] P[Y \in B]$$

2. La *varianza muestral*:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

3. La *desviación típica muestral*:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

Los estadísticos muestrales \bar{X} y S (o bien S^2) miden la tendencia central y la dispersión, respectivamente, de la muestra y, además, son independientes.

Si se conoce la distribución de probabilidad de la población de la que se obtuvo la muestra, entonces muchas veces es posible determinar la distribución de probabilidad de diversos estadísticos muestrales, calculados a partir de la muestra. La distribución de probabilidad de un estadístico muestral se denomina *distribución de muestreo*. A continuación se estudian las distribuciones de muestreo relacionadas con las situaciones más comunes.

Muestreo a Partir de la Distribución Normal

Se analizan las distribuciones de muestreo de estadísticos muestrales relacionados con la distribución Normal:

- Sean X una variable aleatoria con distribución $N(\mu, \sigma)$, y X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n procedente de esta población². Entonces la media muestral tiene distribución $\bar{X} \rightsquigarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

Esta propiedad de la media muestral no es exclusiva para poblaciones Normales, ya que se puede escribir:

$$\left(\frac{\bar{X} - \mu}{\sigma}\right) \sqrt{n} = \frac{\sum_{i=1}^n x_i - n\mu}{\sigma\sqrt{n}}$$

²Esto significa que X_1, X_2, \dots, X_n son variables aleatorias independientes y con distribución de probabilidad $N(\mu, \sigma)$

y por el Teorema Central del Límite, se tiene que la distribución de $\sum_{i=1}^n X_i$ es aproximadamente Normal con media $n\mu$ y varianza $n\sigma^2$, sin importar la distribución de la población. Por lo tanto, la distribución de muestreo de la media muestral es, aproximadamente:

$$\bar{X} \rightsquigarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$$

- Sea X_1, X_2, \dots, X_n una muestra aleatoria procedente de una población $X \rightsquigarrow N(\mu, \sigma)$. Entonces, la variable aleatoria:

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \rightsquigarrow \chi^2(n-1).$$

Por tanto, cuando se realiza el muestreo a partir de la distribución Normal, se pueden calcular probabilidades asociadas a S^2 .

- Sea X_1, X_2, \dots, X_n una muestra aleatoria procedente de una población $X \rightsquigarrow N(\mu, \sigma)$. Entonces, como:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \rightsquigarrow N(0, 1) \quad \text{y} \quad \frac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$$

y además ambos estadísticos son independientes, se tiene que:

$$\frac{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2} \frac{1}{(n-1)}}} = \frac{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{S^2}{\sigma^2}}} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \rightsquigarrow t(n-1).$$

- Dadas $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias simples procedentes de las variables aleatorias independientes $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$, respectivamente, entonces:

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2(n_1-1)}}{\frac{(n_2-1)S_2^2}{\sigma_2^2(n_2-1)}} = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \rightsquigarrow F(n_1-1, n_2-1)$$

donde S_1^2 y S_2^2 son las varianzas muestrales y, además:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \rightsquigarrow N(0, 1).$$

- En las condiciones anteriores, si $\sigma_1^2 = \sigma_2^2$, una estimación de la varianza común es:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Entonces:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow t(n_1 + n_2 - 2).$$

- Si $\sigma_1^2 \neq \sigma_2^2$ y n_1 y n_2 son suficientemente grandes, aproximando mediante el Teorema Central del Límite, se tiene:

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \rightsquigarrow N(0, 1).$$

Muestreo a Partir de una Distribución Binomial

Seguidamente se analizan las distribuciones de muestreo de estadísticos muestrales relacionados con la distribución Binomial.

Sea X una variable aleatoria con distribución de probabilidad $B(1, p)$. Así:

$$P[X = x] = \begin{cases} p & x = 1 \\ 1 - p = q & x = 0 \end{cases},$$

es decir, la variable X toma el valor 1 con probabilidad p , y el valor 0 con probabilidad $1 - p = q$. Al resultado $X = 1$ se le llama éxito y al resultado $X = 0$ se le denomina fracaso. A esta variable aleatoria se le conoce como de *Bernouilli*.

Se extrae una muestra aleatoria de n observaciones, X_1, X_2, \dots, X_n , procedente de la variable aleatoria X , con probabilidad de éxito p constante. Entonces la suma de las observaciones muestrales $X_1 + X_2 + \dots + X_n$ sigue una distribución Binomial con parámetros n y p .

Además, como X_i sólo toma los valores 0 ó 1, la media muestral

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

es una variable aleatoria discreta que puede tomar los valores $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$. Por tanto, es posible obtener la distribución de \bar{X} a partir de la distribución Binomial, ya que

$$P[\bar{X} \leq a] = P\left[\sum_{i=1}^n X_i \leq na\right] = \sum_{k=0}^{[na]} \binom{n}{k} p^k (1-p)^{n-k},$$

donde $[na]$ es el mayor entero menor o igual que na .

Además $\mu_{\bar{X}} = p$ y $\sigma_{\bar{X}}^2 = \frac{p(1-p)}{n}$.

Muestreo a Partir de una Distribución de Poisson

Se considera una muestra aleatoria de tamaño n procedente de una variable aleatoria X con distribución de Poisson con parámetro λ , por ejemplo X_1, X_2, \dots, X_n . La distribución de la suma muestral es:

$$X_1 + X_2 + \dots + X_n \rightsquigarrow P(n\lambda).$$

La distribución de muestreo de la media muestral:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

es una variable aleatoria discreta cuya distribución de probabilidad está dada por:

$$P[\bar{X} \leq a] = P\left[\sum_{i=1}^n X_i \leq na\right] = \sum_{k=0}^{[na]} \frac{e^{-n\lambda} (n\lambda)^k}{k!},$$

donde $[na]$ es el mayor entero menor o igual que na .

La media y la varianza de \bar{X} son $\mu_{\bar{X}} = \lambda$ y $\sigma_{\bar{X}}^2 = \frac{\lambda}{n}$.

Este último resultado se puede extender a dos casos más generales:

1. La suma de n variables aleatorias independientes de Poisson $X_i \sim P(\lambda_i)$ sigue una distribución de Poisson, con parámetro igual a la suma de los parámetros individuales de Poisson:

$$X_1 + X_2 + \cdots + X_n \sim P\left(\sum_{i=1}^n \lambda_i\right).$$

2. Si X_1, X_2, \dots, X_n son variables aleatorias independientes, $X_i \sim P(\lambda_i)$, entonces la combinación lineal:

$$L = a_1X_1 + \cdots + a_nX_n$$

es una variable aleatoria con

$$L \sim P\left(\sum_{i=1}^n \lambda_i a_i\right).$$

Esta propiedad se conoce como *reproductividad*.

Este último caso es bastante útil en situaciones en las que una unidad producida puede tener n diferentes tipos de defectos (cada uno modelado por una distribución de Poisson con parámetro λ_i), y la función que se usa con fines de Control de Calidad es una combinación lineal del número de defectos observados de cada tipo. Podrían escogerse las constantes a_i con $i \in \{1, \dots, n\}$ para dar mayor peso a defectos críticos que a defectos menores. Estos procedimientos se llaman *procedimientos de demérito*.

1.4.2 Estimación de Parámetros

Una variable aleatoria se caracteriza o describe mediante su distribución de probabilidad, la cual es descrita, a su vez, por sus parámetros. En el Control Estadístico de la Calidad se utiliza la distribución de probabilidad para modelizar o describir alguna característica de calidad como, por ejemplo, la longitud de un producto, el porcentaje de defectuosos en un proceso de fabricación, etc. Por lo tanto, interesa realizar inferencias respecto a los parámetros de distribuciones de probabilidad, puesto que suelen ser desconocidos. Para ello, se necesitan procedimientos para estimarlos a partir de los datos muestrales.

Se puede definir el estimador de un parámetro desconocido como un estadístico muestral que corresponde al parámetro. Un valor numérico particular de un estimador, calculado a partir de datos muestrales se llama estimación. Existen dos tipos de estimadores:

1. Un *estimador puntual* es un estadístico muestral que produce un sólo valor numérico como estimación del parámetro desconocido.
2. Un *estimador por intervalo* es un intervalo aleatorio que incluye el valor real del parámetro, con cierto nivel de probabilidad. Estos intervalos aleatorios se denominan normalmente *intervalos de confianza*.

Estimación puntual

Los estimadores puntuales adecuados requieren cierto número de propiedades importantes. Las dos esenciales son:

1. *Insesgadez*: el estimador puntual debe ser insesgado; es decir, el valor esperado del estimador puntual debe ser igual al parámetro estimado.
2. *Mínima varianza*: el estimador debe tener varianza mínima y, como cualquier estimador puntual es una variable aleatoria, un estimador puntual de varianza mínima tendrá una menor varianza que cualquier otro estimador del parámetro.

Se estudian ahora los estimadores puntuales de los parámetros de las distribuciones más usuales:

- Se considera la variable aleatoria X con distribución de probabilidad $f(x)$. Se supone que se desconocen la media μ y la varianza σ^2 de esta distribución. Sus estimadores puntuales son:

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$
$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

La media y la varianza de una variable aleatoria no son necesariamente los parámetros de la distribución. Por ejemplo, los parámetros de la distribución Binomial son n y p y su media y su varianza son np y $np(1-p)$, respectivamente.

- Un estimador puntual adecuado del parámetro λ de una distribución de Poisson es:

$$\hat{\lambda} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

- Un buen estimador puntual del parámetro p de una distribución de Bernoulli es:

$$\hat{p} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

La media muestral y la varianza muestral, \bar{X} y S^2 , son estimadores insesgados de la media y la varianza poblacionales μ y σ^2 , respectivamente; pero S^2 no tiene varianza mínima, aunque se utiliza por ser insesgado.

Estimación por Intervalos

Una estimación por intervalo de un parámetro es el intervalo formado por dos estadísticos muestrales, de manera que el verdadero valor del parámetro se encuentra en dicho intervalo con una cierta probabilidad. Para elaborar un estimador por intervalo para un parámetro θ , es necesario encontrar dos estadísticos w_1 y w_2 , tales que:

$$P[w_1 \leq \theta \leq w_2] = 1 - \alpha$$

El intervalo resultante, (w_1, w_2) , se denomina *intervalo de confianza* al $100(1 - \alpha)\%$ para el parámetro desconocido. w_1 y w_2 son, respectivamente, los límites inferior y superior de confianza; $1 - \alpha$ se llama nivel de confianza y α es el nivel de significación o margen de error.

El significado del intervalo de confianza es que si se determina un gran número de intervalos de confianza para un determinado parámetro procedente de una variable aleatoria, cada uno a partir de una muestra aleatoria

distinta, entonces el $100(1 - \alpha)\%$ de estos intervalos contendrán el verdadero valor del parámetro θ . Por tanto, los intervalos de confianza tienen una interpretación frecuencial.

El intervalo $w_1 \leq \theta \leq w_2$ es un intervalo de confianza *bilateral*, ya que especifica tanto un límite inferior como uno superior para θ . A veces, en algunas aplicaciones del Control Estadístico de la Calidad, es conveniente utilizar un intervalo de confianza *unilateral*, el cual puede ser de dos tipos:

1. Inferior al $100(1 - \alpha)\%$ para θ estaría dado por (w_1, ∞) , donde w_1 es el límite inferior de confianza y se escoge de manera que $P[w_1 \leq \theta] = 1 - \alpha$.
2. Superior al $100(1 - \alpha)\%$ para θ correspondería a $(-\infty, w_2)$, donde $P[\theta \leq w_2] = 1 - \alpha$ y w_2 es el límite superior de confianza.

Se presentan algunos de los intervalos de confianza más usuales, que se obtienen a partir de las distribuciones de muestreo de los estadísticos muestrales estudiados.

(a) Intervalo de confianza para la media con varianza conocida: Sean X una variable aleatoria con varianza conocida σ^2 y X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n . Entonces:

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\bar{X} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

donde z_α es el cuantil de orden α , esto es, el valor que deja por debajo de él una probabilidad igual a α de la distribución Normal de parámetros 0 y 1:

$$P[Z \leq z_\alpha] = \alpha \quad \text{con } Z \sim N(0, 1)$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\mu \leq \bar{X} + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\bar{X} - z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \leq \mu$$

Estos intervalos no son exactos, salvo que la población sea Normal, sino que son aproximaciones debidas al Teorema Central del Límite, como ya se vio al estudiar la distribución de muestreo.

(b) Intervalo de confianza para la media de una distribución Normal con varianza desconocida: Sea X una variable aleatoria con distribución $N(\mu, \sigma)$ con μ y σ desconocidas, y sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n .

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\bar{X} - t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\mu \leq \bar{X} + t_{n-1, 1-\alpha} \frac{S}{\sqrt{n}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\bar{X} - t_{n-1, 1-\alpha} \frac{S}{\sqrt{n}} \leq \mu$$

(c) Intervalo de confianza para la varianza de una distribución Normal: Sea X una variable aleatoria con distribución $N(\mu, \sigma)$ con σ desconocida, y sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n .

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\frac{(n-1)S^2}{\chi_{n-1, 1-\frac{\alpha}{2}}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{n-1, \frac{\alpha}{2}}^2}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\sigma^2 \leq \frac{(n-1)S^2}{\chi_{n-1, \alpha}^2}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\frac{(n-1)S^2}{\chi_{n-1,1-\alpha}^2} \leq \sigma^2$$

(d) Intervalo de confianza para la diferencia de medias con varianzas

conocidas: Sean X_1, X_2 variables aleatorias independientes con medias μ_1 y μ_2 , y varianzas conocidas σ_1^2 y σ_2^2 , respectivamente. Sean $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias de X_1 y X_2 .

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$(\bar{X}_1 - \bar{X}_2) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z_{1-\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$(\bar{X}_1 - \bar{X}_2) - z_{1-\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2$$

(e) Intervalo de confianza para la diferencia de medias de dos distribuciones Normales con varianzas desconocidas pero iguales:

Sean las variables aleatorias independientes $X_1 \sim N(\mu_1, \sigma_1)$ y $X_2 \sim N(\mu_2, \sigma_2)$ con $\sigma_1^2 = \sigma_2^2$ pero desconocidas. Sean $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias de X_1 y X_2 , respectivamente.

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\begin{aligned} (\bar{X}_1 - \bar{X}_2) - t_{n_1+n_2-2, 1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} &\leq \mu_1 - \mu_2 \leq \\ &\leq (\bar{X}_1 - \bar{X}_2) + t_{n_1+n_2-2, 1-\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \end{aligned}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + t_{n_1+n_2-2, 1-\alpha} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$(\bar{X}_1 - \bar{X}_2) - t_{n_1+n_2-2, 1-\alpha} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2$$

(f) Intervalo de confianza para la diferencia de medias de dos distribuciones Normales con varianzas desconocidas pero distintas: Sean las variables aleatorias independientes $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ con σ_1^2 y σ_2^2 desconocidas y distintas. Sean $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias grandes de X_1 y X_2 respectivamente. Aproximando mediante el Teorema Central del Límite, se tiene:

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$(\bar{X}_1 - \bar{X}_2) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z_{1-\alpha} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$(\bar{X}_1 - \bar{X}_2) - z_{1-\alpha} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \leq \mu_1 - \mu_2$$

(g) Intervalos de confianza para el cociente de varianzas de dos distribuciones Normales: Sean las variables aleatorias independientes $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$. Sean $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias de X_1 y X_2 respectivamente.

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\frac{S_1^2}{S_2^2} f_{n_2-1, n_1-1, \frac{\alpha}{2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} f_{n_2-1, n_1-1, 1-\frac{\alpha}{2}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} f_{n_2-1, n_1-1, 1-\alpha}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\frac{S_1^2}{S_2^2} f_{n_2-1, n_1-1, \alpha} \leq \frac{\sigma_1^2}{\sigma_2^2}$$

(h) Intervalo de confianza para proporciones:

1. *Para una población o proporción:* Sea X una variable aleatoria con distribución $B(1, p)$. Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n . Se pretenden construir intervalos de confianza para el parámetro desconocido p . Si se aproxima, utilizando el Teorema Central del Límite, por la distribución Normal:

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$p \leq \hat{p} + z_{1-\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\hat{p} - z_{1-\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p$$

2. *Para dos poblaciones o proporciones:* Sean las variables aleatorias independientes $X_1 \rightsquigarrow B(1, p_1)$ y $X_2 \rightsquigarrow B(1, p_2)$ y $X_{11}, X_{12}, \dots, X_{1n_1}$ y $X_{21}, X_{22}, \dots, X_{2n_2}$ muestras aleatorias procedentes de X_1 y X_2 respectivamente. Se pretende construir intervalos de confianza para la diferencia de los parámetros desconocidos $p_1 - p_2$. Si se aproxima por la distribución Normal:

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\begin{aligned} \hat{p}_1 - \hat{p}_2 - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} &\leq p_1 - p_2 \leq \\ &\leq \hat{p}_1 - \hat{p}_2 + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \end{aligned}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + z_{1-\alpha} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\hat{p}_1 - \hat{p}_2 - z_{1-\alpha} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \leq p_1 - p_2$$

(i) Intervalo de confianza para el parámetro de una Poisson: Sea X una variable aleatoria con distribución $P(\lambda)$ y sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n . Si se aproxima, utilizando el Teorema Central del Límite, por la distribución Normal:

- Intervalo de confianza bilateral al nivel de confianza $1 - \alpha$:

$$\bar{X} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}} \leq \lambda \leq \bar{X} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}}{n}}$$

- Intervalo de confianza superior al nivel de confianza $1 - \alpha$:

$$\lambda \leq \bar{X} + z_{1-\alpha} \sqrt{\frac{\bar{X}}{n}}$$

- Intervalo de confianza inferior al nivel de confianza $1 - \alpha$:

$$\bar{X} - z_{1-\alpha} \sqrt{\frac{\bar{X}}{n}} \leq \lambda$$

1.4.3 Contraste de hipótesis

Un *contraste de hipótesis* es un procedimiento mediante el cual se decide si una afirmación sobre los valores de los parámetros de una distribución de probabilidad se puede aceptar o no. Un ejemplo puede ser estudiar si, con base en la información muestral, la proporción de productos defectuosos es mayor que la máxima aceptable.

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n procedente de una variable aleatoria X , con distribución de probabilidad $F(x; \theta)$, donde el parámetro θ toma valores dentro del espacio paramétrico Θ .

Se llama *hipótesis nula* (H_0) a una afirmación relativa al parámetro θ desconocido, $H_0 : \theta \in \Theta_0$, es decir, la hipótesis nula consiste en afirmar que el parámetro toma determinados valores del espacio paramétrico Θ ($\Theta_0 \subset \Theta$). La hipótesis nula siempre contendrá el signo igual y al valor del parámetro aceptado previamente como correcto.

Se llama *hipótesis alternativa* (H_1) a otra hipótesis relativa al parámetro desconocido, que por lo general suele ser la hipótesis opuesta a la nula $H_1 : \theta \in \Theta_1$ (en general, $\Theta_1 = \Theta - \Theta_0$, aunque no siempre). Con ella, se pretende ver si el verdadero valor del parámetro ha cambiado respecto del aceptado previamente en el sentido considerado en ella.

Dependiendo de que la hipótesis alternativa esté a uno o a ambos lados de la hipótesis nula, se hablará de *hipótesis alternativa unilateral o bilateral*.

Los procedimientos de contraste de hipótesis son muy útiles en muchos tipos de problemas de Control Estadístico de Calidad y son la base de la mayoría de las técnicas de Control Estadístico de Procesos. Una parte importante de cualquier problema de contraste de hipótesis es determinar los valores paramétricos especificados en las hipótesis nula y alternativa. En general, esto se realiza de una de las tres maneras siguientes:

1. Los valores pueden resultar de pruebas o conocimientos anteriores. Esto sucede con frecuencia en el Control Estadístico de Calidad, donde se usa información anterior para especificar los valores de un parámetro correspondiente a un estado de control, y después se prueba periódicamente la hipótesis de que el valor del parámetro no ha cambiado.
2. Los valores pueden ser el resultado de alguna teoría o modelo del proceso.

3. Los valores escogidos para el parámetro pueden ser el resultado de especificaciones de diseño. Pueden utilizarse procedimientos estadísticos de contraste de hipótesis para verificar la conformidad de los parámetros del proceso con sus valores especificados, o para ayudar a modificar un proceso hasta obtener valores requeridos.

Para realizar un contraste de hipótesis, se toma una muestra aleatoria de la población observada, se calcula un estadístico de contraste apropiado, y después, a partir de los resultados obtenidos en la muestra, se decide rechazar o no la hipótesis nula H_0 . El conjunto de valores del estadístico de contraste que lleva al rechazo de H_0 se llama *región crítica o región de rechazo* del contraste.

Se pueden cometer dos tipos de errores al realizar un contraste de hipótesis:

1. Rechazar H_0 cuando es verdadera, cometiendo un *error de tipo I*.
2. No rechazar H_0 cuando es falsa, cometiendo entonces un *error de tipo II*.

Las probabilidades de que ocurran estos dos tipos de errores se denotan como:

$$\begin{aligned}\alpha &= P[\text{error de tipo I}] = P[\text{rechazar } H_0 \mid H_0 \text{ es verdadera}] \\ \beta &= P[\text{error de tipo II}] = P[\text{no rechazar } H_0 \mid H_0 \text{ es falsa}]\end{aligned}$$

Algunas veces conviene más trabajar con la *potencia del contraste*, definida como:

$$\text{Potencia} = 1 - \beta = P[\text{rechazar } H_0 \mid H_0 \text{ es falsa}]$$

Así, la potencia del contraste es la probabilidad de rechazar correctamente H_0 . En el Control de Calidad, α se llama a veces riesgo del fabricante, ya que es la probabilidad de que sea rechazado un lote aceptable, o la probabilidad de que sea rechazado por funcionar insatisfactoriamente un proceso que produce valores aceptables de una característica de calidad en particular. Así mismo, β se denomina a veces riesgo del consumidor, ya que es la probabilidad de aceptar un lote de calidad deficiente, o la probabilidad

de que se deje seguir operando un proceso que trabaja insatisfactoriamente respecto a alguna característica de calidad.

Lo ideal sería encontrar una región crítica en la que ambos errores fueran nulos, sin embargo, para un tamaño muestral fijo, si se trata de disminuir un error, automáticamente aumenta el otro. Por ello, el procedimiento general para una prueba de hipótesis es especificar un valor de α , también llamado *nivel de significación*, y después diseñar un procedimiento para obtener el valor más pequeño posible de β . El riesgo β es por lo general una función del tamaño muestral, y se controla indirectamente. Cuanto más grande sea el tamaño muestral para la prueba, menor será el riesgo β , para un valor de α fijo.

Un contraste de hipótesis paramétrico se resuelve siguiendo los siguientes pasos:

1. Lo primero que se debe realizar es obtener la hipótesis nula y la hipótesis alternativa del problema. La hipótesis nula deberá contener el signo igual.
2. Se estudian las condiciones del problema para elegir el estadístico h de contraste adecuado, a elegir entre los distintos estimadores muestrales estudiados, donde cada uno sigue una determinada distribución de probabilidad H .
3. Se calcula la realización del estadístico muestral h_{exp} , supuesta cierta la hipótesis nula; es decir, se sustituyen todos los valores muestrales y se obtiene el *valor experimental del contraste*.
4. Se compara el valor experimental obtenido con los valores teóricos de la distribución del estadístico muestral utilizado, dependiendo del tipo de contraste:

- Contraste bilateral:

$$\left. \begin{array}{l} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{array} \right\}$$

Si $h_{\text{exp}} < H_{\frac{\alpha}{2}}$ ó $h_{\text{exp}} > H_{1-\frac{\alpha}{2}}$, se rechaza la hipótesis nula al nivel de significación α .

- Contraste Unilateral:

$$\left. \begin{array}{l} H_0 : \theta \geq \theta_0 \\ H_1 : \theta < \theta_0 \end{array} \right\}$$

Si $h_{\text{exp}} < H_\alpha$ hay evidencia suficiente para rechazar la hipótesis nula al nivel de significación α .

- Contraste Unilateral:

$$\left. \begin{array}{l} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{array} \right\}$$

Si $h_{\text{exp}} > H_{1-\alpha}$, hay evidencia suficiente para rechazar la hipótesis nula al nivel de significación α .

Ejemplo 22 *Se quiere contrastar la hipótesis, con un nivel de significación del 5%, de que el peso medio de un determinado producto es de 180 gr. Supuesta la hipótesis de normalidad del peso del producto, se toma una muestra de 15 productos, los cuales se pesan y se obtiene una media muestral de 185.87 gr. y una varianza muestral de 104.9810 gr².*

Entonces:

$$\left. \begin{array}{l} H_0 : \mu = 180 \\ H_1 : \mu \neq 180 \end{array} \right\}$$

Se está en el caso de contraste de hipótesis para la media con varianza poblacional desconocida, por lo que la realización del estadístico muestral es:

$$t_{\text{exp}} = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{185.87 - 180}{\frac{\sqrt{104.9810}}{\sqrt{15}}} = 2.2189 \rightsquigarrow t(15 - 1) = t(14)$$

Entonces, como $t_{\text{exp}} = 2.2189 \notin (t_{14,0.025}, t_{14,0.975}) = (-2.14, 2.14)$, se rechaza la hipótesis nula con nivel de significación del 5%, es decir, no se puede aceptar que la longitud media es de 180 gr.

Probabilidad de error de tipo II

En la mayoría de los casos de contraste de hipótesis, es importante determinar la probabilidad del error de tipo II asociado a la prueba. De

manera equivalente, se puede elegir evaluar la potencia del contraste. Para ilustrar cómo se calcula, se obtiene la probabilidad de cometer error de tipo II asociado al siguiente contraste:

$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right\}$$

con varianza conocida σ^2 .

El estadístico de contraste para esta hipótesis es:

$$Z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

y bajo H_0 , $Z_0 \rightsquigarrow N(0, 1)$.

Para calcular β se supone que H_0 es falsa y se calcula la distribución de Z_0 . Se supone que la media de la distribución es realmente $\mu_0 + \delta = \mu_1$, donde $\delta > 0$. Por lo tanto, la hipótesis alternativa $H_1 : \mu \neq \mu_0$ es verdadera. Bajo esta hipótesis:

$$E[Z_0] = E\left[\frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right] = \frac{1}{\frac{\sigma}{\sqrt{n}}} (E[\bar{X}] - \mu_0) = \frac{1}{\frac{\sigma}{\sqrt{n}}} (\mu_0 + \delta - \mu_0) = \frac{\delta\sqrt{n}}{\sigma}$$

Luego

$$Z_0 \rightsquigarrow N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$$

Así β es la probabilidad de que Z_0 caiga entre $z_{\frac{\alpha}{2}}$ y $z_{1-\frac{\alpha}{2}}$, dado que H_1 es verdadera, esto es:

$$\begin{aligned} P\left[z_{\frac{\alpha}{2}} \leq Z_0 \leq z_{1-\frac{\alpha}{2}}\right] &= P\left[Z_0 \leq z_{1-\frac{\alpha}{2}}\right] - P\left[Z_0 \leq z_{\frac{\alpha}{2}}\right] = \\ &= P\left[Z_0 - \frac{\delta\sqrt{n}}{\sigma} \leq z_{1-\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right] - P\left[Z_0 - \frac{\delta\sqrt{n}}{\sigma} \leq z_{\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right] = \\ &= P\left[Z \leq z_{1-\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right] - P\left[Z \leq z_{\frac{\alpha}{2}} - \frac{\delta\sqrt{n}}{\sigma}\right] = \beta \end{aligned}$$

β es una función de n , δ y α . Se acostumbra a trazar gráficas que representan la relación entre estos parámetros. Estas gráficas se suelen denominar *curvas características operativas* o *curvas características de operación* (OC).

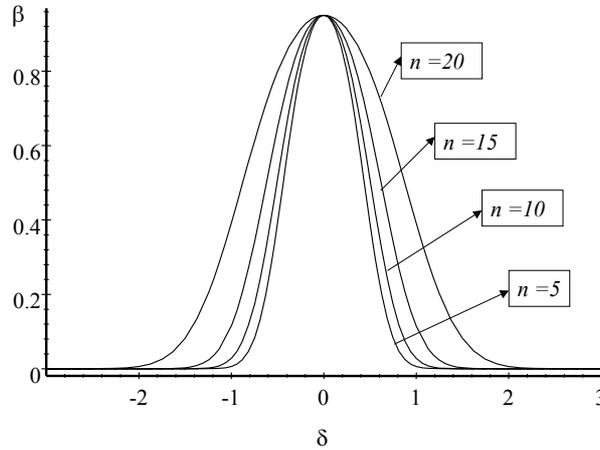


Figura 1.6: Curva OC para distintos valores de n

El parámetro en el eje vertical de estas curvas es β , y el parámetro en el eje horizontal es $d = \frac{|\delta|}{\sigma}$.

En la Figura 1.6 se pueden ver distintas curvas OC para varios valores de n .

Examinando las curvas características operativas, se observa que:

1. Cuanto más se aleje la verdadera media μ_1 del valor de la hipótesis nula μ_0 (es decir, cuanto mayor es el valor de δ), tanto más pequeña será la probabilidad de error de tipo II, para un valor dado de n y de α . O sea, para un tamaño muestral y un valor de α especificados, la prueba detectará con mayor facilidad diferencias grandes que pequeñas.
2. Al aumentar el tamaño muestral n , la probabilidad de error de tipo II decrece para α y δ especificados. Es decir, para detectar una diferencia en particular se puede hacer la prueba más poderosa aumentando el tamaño muestral.

1.5 Ejercicios

1. Los siguientes datos son los volúmenes en ml. de llenado de 16 botellas que se seleccionaron de un proceso de llenado para estudiar el volumen promedio: 506, 508, 499, 503, 504, 510, 497, 512, 514, 505, 493, 496, 506, 502, 509, 496. Si el volumen de llenado de cada botella es una variable aleatoria Normal con una desviación típica $\sigma = 5$ ml., obtenga los intervalos de confianza estimados del 90, 95 y 99%, para la media de llenado de este proceso.
2. Una muestra aleatoria de los salarios (en euros) por hora para nueve trabajadores es: 10.5, 11, 9.5, 12, 10, 11.5, 13, 9, 8.5. Si el muestreo se realizó sobre una población distribuida Normal, construya los intervalos de confianza estimados del 95% para los salarios por hora promedio para todos los trabajadores.
3. El precio de un determinado producto en distintas tiendas es distinto. Se seleccionan, en forma aleatoria, 12 tiendas y se observa el precio del producto, siendo 3.01, 3.05, 2.99, 2.99, 3.00, 3.02, 2.98, 2.99, 2.97, 2.97, 3.02, 3.01. Suponiendo que el precio de este producto es una variable aleatoria normalmente distribuida, determine un intervalo de confianza del 99% para la varianza σ^2 .
4. Dos hospitales desean comparar el tiempo promedio de espera (en días) para que sus pacientes sean operados. En cada hospital se anotaron los tiempos de espera de 100 pacientes seleccionados al azar. Las medias muestrales son las siguientes:

$$\bar{x}_1 = 50.2 \quad \bar{x}_2 = 52.9$$

Si se supone que el muestreo se llevó a cabo sobre dos poblaciones distribuidas Normales e independientes con desviaciones típicas

$$\sigma_1 = 10 \quad \sigma_2 = 12$$

obtenga los intervalos de confianza estimados del 95 y 99% para la diferencia entre las medias del tiempo de espera.

5. Se pretende comparar el rendimiento de dos máquinas que realizan el mismo trabajo. El comprador está interesado en estimar la verdadera

diferencia entre ellas. Para cada máquina se seleccionan 12 artículos y se estudian los resultados, que aparecen en la siguiente tabla:

Máquina A	Máquina B
428	462
419	448
458	435
439	465
441	429
456	472
463	453
429	459
438	427
445	468
441	452
463	447

Si el muestreo se llevó a cabo sobre las dos distribuciones Normales e independientes con varianzas desconocidas, obtenga los intervalos de confianza estimados del 99% para $\mu_S - \mu_N$ suponiendo varianzas iguales y varianzas distintas.

6. Recientemente se han obtenido muestras aleatorias de los salarios de los trabajadores que han estudiado la I.T. Informática y los que han estudiado I.T. Industrial. Los datos son los siguientes:

	n	\bar{x}	S_X^2
Informática	10	16250	1187222.22
Industrial	14	15400	1352307.69

Obtenga un intervalo de confianza para el cociente de varianzas $\frac{\sigma_A^2}{\sigma_E^2}$ con un margen de error del 10%.

7. Un fabricante asegura que el porcentaje de ordenadores defectuosos es del 5%. El distribuidor decide comprobar la afirmación del fabricante seleccionando 200 ordenadores al azar y probándolos. ¿Deberá sospechar el distribuidor de la afirmación del fabricante si se descubren un total de 19 unidades defectuosas en la muestra? Utilice un nivel de confianza del 99%.

8. Un médico desea estimar la diferencia entre la proporción de hombres y mujeres, en edad madura, que fuman en exceso y que desarrollan un cáncer pulmonar en los siguientes cinco años. Para ello selecciona dos muestras, una de hombres y otra de mujeres verificando las condiciones anteriores. Los datos son los siguientes:

	Nº fumadores	Enfermos de cáncer
Hombres	100	85
Mujeres	110	60

Calcule un intervalo de confianza al 95% para la diferencia entre las proporciones de enfermos. ¿Cuáles son los límites para el intervalo de confianza a un nivel de confianza del 99%?

9. Los siguientes datos representan los tiempos de atención (en minutos) de una consulta médica para 20 pacientes seleccionados aleatoriamente: 9.8, 10.4, 10.6, 9.6, 9.7, 9.9, 10.9, 11.1, 9.6, 10.2, 10.3, 9.6, 9.9, 11.2, 10.6, 9.8, 10.5, 10.1, 10.5, 9.7. Si el tiempo de atención es una variable aleatoria Normal con media μ y desviación típica $\sigma = 0.6$ minutos, ¿existe alguna razón para creer a un nivel de 0.05, que el tiempo de atención promedio es mayor de 10 minutos?
10. El tiempo promedio de montaje de un artículo es de 5 horas. Para comprobarlo, se toma una muestra de 11 artículos y obtienen los siguientes resultados: 4.8, 5.6, 5.3, 5.2, 4.9, 4.7, 5.7, 4.9, 5.7, 4.9, 4.6. Si se supone que el tiempo de montaje se encuentra modelado en forma adecuada por una distribución Normal, ¿es cierto lo afirmado con anterioridad con $\alpha = 0.02$?
11. En un proceso, la desviación típica en el volumen debe ser de dos litros. Los volúmenes de 25 artículos seleccionados al azar dieron como resultado una desviación típica muestral de 2.8 litros. Si los volúmenes se encuentran Normalmente distribuidos, determine si la varianza de éstos es diferente del valor necesario. Empléese $\alpha = 0.02$.
12. Se pretende comparar el tiempo de funcionamiento de las bombillas de dos marcas distintas para descubrir posibles diferencias significativas entre ellas. Dichos tiempos de funcionamiento se distribuyen normalmente con varianzas 10000 y 12100, respectivamente. En idénticas

condiciones, se estudia el tiempo de funcionamiento de 8 bombillas de cada marca:

Marca A	Marca B
1010	1040
980	1000
880	870
900	965
1205	1185
1060	1030
870	860
990	990

Para un nivel de significación $\alpha = 0.025$ verifique la afirmación de que existe una diferencia significativa.

13. Se desean comparar longitudes de tornillos de dos proveedores distintos. Dichas longitudes están distribuidas Normalmente en ambos proveedores, pero se desconocen sus varianzas, aunque se sabe que son iguales. Los datos son los siguientes:

	Empresa	
	A	B
Media	3.45	3.51
S^2	0.005	0.004
n	40	35

¿Hay diferencia entre las medias? (Asuma que $\alpha = 0.05$).

14. Un inversor quiere analizar los riesgos asociados con dos diferentes mercados, A y B. El riesgo de un mercado dado se mide por la variación en los cambios diarios de precios. El inversor piensa que el riesgo asociado con el mercado B es mayor que el del mercado A. Se obtienen muestras aleatorias de 21 cambios de precio diarios para el mercado A y de 16 para el mercado B. Se obtienen los siguientes resultados:

Mercado A	Mercado B
$\bar{x}_A = 0.3$	$\bar{x}_B = 0.4$
$S_A = 0.25$	$S_B = 0.45$

Si se supone que las muestras provienen de dos poblaciones Normales e independientes a un nivel de $\alpha = 0.05$, ¿encuentra apoyo la creencia del inversor?

Capítulo 2

Control Estadístico de Procesos

2.1 Introducción Histórica

Desde sus orígenes, el hombre se ha preocupado por la calidad de los productos, si bien el concepto de calidad ha sufrido numerosos cambios a lo largo de la historia.

La aparición de las primeras ciudades trajo como consecuencia la creación del mercado para bienes y servicios, lo que permitió el desarrollo inicial de especificaciones para productos y procesos, dando como resultado nuevas formas de organización.

En la actividad mercantil, entre los siglos XII y XVIII, el concepto de Calidad siempre está implícito, si bien en principio no se disponía de especificaciones técnicas. La finalidad del artesano era satisfacer al cliente, a sí mismo por un trabajo bien hecho y crear un producto único, independientemente del coste o esfuerzo necesario para ello. Con un sistema de producción artesanal los problemas de Calidad pueden resolverse con relativa facilidad puesto que el fabricante, el comprador y las mercancías están presentes simultáneamente. Cuando la actividad artesanal se masifica, algunos gremios imponen unos reglamentos que normalizaban la Calidad de los productos acabados. Algunas mercancías eran inspeccionadas y selladas por el gremio.

A finales del siglo XIX, con la revolución industrial, comienza la producción en serie para satisfacer una gran demanda de bienes y obtener

beneficios, es decir, se identifica producción con calidad. La producción industrial reduce los costos de producción respecto al sistema artesanal, pero masifica los centros de trabajo por lo que surge la necesidad de establecer una división del trabajo. De esta forma, el operario que ejecuta el trabajo no tiene preocupación por el resultado del mismo; otro operario al final de la cadena inspeccionará la producción y se encargará de seleccionar adecuadamente los productos.

La primera guerra mundial pone al descubierto que el armamento alemán, por ser más uniforme y utilizar unas tolerancias de fabricación más estrictas, es más eficaz que el del resto de países.

Durante el periodo entre guerras se introducen técnicas estadísticas en la inspección final, evitando tener que controlar todos los productos. Posteriormente se establecen controles intermedios del producto durante el proceso, con lo que se evita rechazar un producto totalmente acabado.

Las *técnicas estadísticas* de Control de Calidad tienen su comienzo en 1924 cuando *Walter A. Shewhart*, considerado como el padre del Control Estadístico de la Calidad, desarrolló el concepto estadístico de los Diagramas o Gráficos de Control. Su gran trabajo titulado *Control Económico de Calidad de Productos Manufacturados (Economic Control of Quality Manufactured Product)*, publicado en 1931, establece los principios básicos del Control de Calidad Moderno (véase Shewart (1931)).

Walter Shewhart se graduó en la Universidad de Illinois y recibió el Doctorado en Física en la Universidad de California, en Berkeley, en 1917. Impartió clases en las Universidades de Illinois y California y dirigió el Departamento de Física de la *Escuela Normal Wisconsin* en LaCrosse. Desarrolló la mayor parte de su carrera profesional como ingeniero en la compañía *Western Electric* desde 1918 hasta 1924 y posteriormente, hasta su jubilación en 1956, en los *Laboratorios Bell Telephone*, donde desempeñó diversas tareas como miembro del equipo técnico. Asimismo dio conferencias de Control de Calidad y estadística aplicada en la Universidad de Londres e incluso en India. Entre otras, fue miembro de la *Sociedad Estadística Real Inglesa* y de las *Asociaciones de Estadística Americana* y de *Calcula*. Durante más de veinte años fue el primer editor de las *Series de Estadística Matemática* publicadas por la editorial *John Wiley and Sons*.

En 1939, Shewhart escribió *Métodos Estadísticos desde el punto de vista del Control de Calidad*, lo que supuso su gran reconocimiento en la comunidad estadística. Uno de sus numerosos artículos fue el memorándum del

16 de Mayo de 1924, en el que propuso el uso de los Gráficos de Control.

En la década de los años 20, *Harold F. Dodge* y *Harold G. Roming* desarrollaron el muestreo para aceptación como alternativa a la inspección al cien por cien. Pero, sin duda, es después de la Segunda Guerra Mundial cuando las técnicas de Control de Calidad empiezan a ser utilizadas de forma general.

Cuando Japón atacó Pearl Harbour, los Estados Unidos decidieron participar activamente en la II Guerra Mundial, de modo que la demanda de material bélico se incrementó considerablemente, así como la necesidad de perfeccionar el armamento. Por tanto, se planteó la necesidad de educar a los trabajadores en el control estadístico de la calidad de los productos fabricados.

Tras esta guerra, Japón se enfrentaba a la ardua tarea de reconstruir su país. En aquel momento, las fuerzas de ocupación de los E.E.U.U. decidieron apoyar en la reconstrucción de la economía y la infraestructura de forma directa, con el objetivo de evitar que el país nipón recuperase su capacidad bélica. Así pues, enviaron gran parte de expertos estadounidenses para ayudar en esta labor que, además, necesitaban ganarse previamente la confianza de los japoneses que aún los veían como enemigos. Para ello crearon la *Sección de Comunicación Civil (CCS)* cuya misión era difundir mensajes pro-E.E.U.U. en la población utilizando, entre otros, programas de radio. Sin embargo, la población no disponía de aparatos de radio; por consiguiente, se establecieron infraestructuras para la fabricación de radios cuyos administradores más experimentados fueron apartados durante la guerra, de modo que el personal con el que se contaba carecía de formación y experiencia con una consecuente disminución de la calidad de los productos. Para paliar este problema se creó el *Laboratorio de Pruebas Eléctrico Nacional (NETL)* cuya misión era controlar la calidad. Sin embargo esta estrategia no obtuvo buenos resultados a largo plazo por lo que la CCS, en colaboración con la *Unión de Científicos Ingenieros de Japón (JUSE)*, decidió formar a la nueva generación de administradores. Entre las actividades de formación se incluyó el Control Estadístico de la Calidad (*Statistical Quality Control, SQC*) junto con las aportaciones en este campo de Walter Shewhart. Para la JUSE el Control de Calidad fue decisivo para la victoria de los E.E.U.U. en la guerra, así que centró su interés en este campo solicitando a la CCS que les recomendase expertos para profundizar en el tema. En aquel entonces Walter Shewhart no estaba disponible, así

que les recomendaron al profesor *Edwards Deming* de la Universidad de Columbia que había estudiado y aplicado los métodos del primero (algunos autores consideran a Deming como el padre del control de la calidad, si bien los estudios de éste se basaron en los de Shewhart). Así pues, en 1950 y durante dos meses, Deming entrenó a cientos de ingenieros, administradores y ejecutivos de primer nivel.

El 16 de Febrero de 1946, diecisiete sociedades de Control de Calidad formaron la *Sociedad Americana para el Control de Calidad (ASQC)*. La primera reunión tuvo lugar en Junio de 1947 en Chicago y tuvo tanto éxito que se convirtió en un acontecimiento anual. Fue en la segunda reunión cuando se establecieron las bases y el código ético. En 1997 la ASQC pasó a ser la *Sociedad Americana para la Calidad (ASQ)*. Esta sociedad promueve el uso de técnicas de Control de Calidad para todo tipo de productos y servicios, además de ofrecer conferencias, publicaciones técnicas y programas de adiestramiento para asegurar la calidad.

Más adelante, en los años 60, *Philip B. Crosby* propuso un programa de 14 pasos a los que denominó “cero defectos”, a través de los cuales hizo entender a los directivos que cuando se exige perfección ésta puede lograrse, pero para ello la alta gerencia tiene que motivar a sus trabajadores, así como utilizar la prevención para reducir errores controlables (véase Crosby (1979)).

En los años 80, *Genichi Taguchi* introdujo el *Método de Taguchi* (Taguchi (1986)) donde combina la estadística y la ingeniería para la reducción de costos y mejora de la calidad en el diseño de productos y los procesos de fabricación. En sus métodos emplea la experimentación a pequeña escala con la finalidad de reducir la variación y descubrir diseños robustos y baratos para la fabricación en serie. Las aplicaciones más avanzadas del Método Taguchi, permiten desarrollar tecnología flexible para el diseño y fabricación de familias de productos de alta calidad, reduciendo los tiempos de investigación, desarrollo y entrega del diseño.

En la década de los 90, *Armand Feigenbaum* introdujo el concepto de Control de Calidad Total, basado en el enfoque total de sistemas (véase Feigenbaum (1991)). Todas las áreas como finanzas, ventas y mercadeo, compras y otras, eran responsables de la calidad. Para Feigenbaum, era importante que las compañías contaran con procesos bien planeados y documentados para:

- El control de diseños nuevos.
- El control de adquisición de materiales.
- El control del producto.
- La realización de estudios especiales del proceso.

Feigenbaum define el Control Total de Calidad como “un sistema efectivo para la integración de los esfuerzos de desarrollo, mantenimiento y mejoramiento que los diferentes grupos de una organización realizan para poder proporcionar un producto o servicio en los niveles más económicos para satisfacción de las necesidades del usuario.”

En cualquier caso, notamos que el Control de Calidad está en constante desarrollo, no sólo en los métodos, sino también en la filosofía, esto es, en la forma de entender la calidad.

2.2 ¿En qué consiste el Control de Calidad?

Son diversas las definiciones que los distintos autores dan del concepto de calidad. En general, puede decirse que un artículo tiene *calidad* cuando es apto para su utilización y responde a las necesidades del consumidor. Con el *Control de Calidad* se intenta asegurar la calidad de los productos fabricados (coches, electrodomésticos, muebles, ordenadores, maquinaria...) o de los servicios (transporte, asistencia médica, educación...). En consecuencia, para conseguir que un producto sea de alta calidad se ha de diseñar de modo que se recojan las necesidades del consumidor; asimismo, es necesario diseñar el proceso que permita fabricar dicho artículo.

El artículo se diseña según un nivel de calidad. Las variaciones en los grados o niveles de calidad son intencionales, en el siguiente sentido: todos los automóviles tienen la misma finalidad - proporcionar al usuario un transporte seguro - sin embargo, difieren en tamaño, equipo, presentación, rendimiento... Además, el resultado final depende del proceso de fabricación, y se pretende que éste se aproxime lo máximo posible al modelo ideado. Para determinar lo que se acerca el artículo fabricado al modelo buscado, se miden las denominadas *características de calidad* como, por ejemplo, la longitud, el diámetro, el peso, el color, la resistencia... De este modo, si el artículo fabricado se aleja del modelo más de lo deseado, se

rechaza. El valor a partir del cual rechazamos el producto se llama *nivel de tolerancia*; así, si estamos fabricando tornillos de 1 cm de longitud, la característica de calidad que mediremos será la longitud, y si decimos que el nivel de tolerancia es de ± 0.05 cm, rechazaremos todos los tornillos que midan menos de 0.95 cm y más de 1.05 cm. Por tanto, cuanto menor sea el nivel de tolerancia, mayor será la calidad del producto. Es claro que el interés del fabricante reside en que la mayoría de los productos estén dentro del nivel de tolerancia, ya que no interesa fabricar productos que tengan que ser rechazados.

En definitiva, el *Control de Calidad* es el conjunto de técnicas y procedimientos mediante los cuales se miden las características de calidad de un producto, se comparan con especificaciones y se toman acciones correctivas apropiadas cuando hay muchos elementos que no cumplen los requisitos.

Las *técnicas básicas* empleadas en el Control de Calidad se citan a continuación:

- *Diseño de Experimentos*, para las etapas de diseño del artículo y en las primeras etapas de fabricación.
- *Control estadístico de procesos*, para regular y vigilar el proceso.
- *Muestreo para la aceptación*, cuando el objetivo de la inspección es la aceptación o el rechazo de un producto, con base en la conformidad respecto a un estándar.

2.3 Algunas Definiciones en Control de Calidad

La producción de un artículo se sustenta en un proyecto de fabricación que ha de ser de calidad. La *calidad de proyecto* de un artículo está estrechamente relacionada con la precisión de las especificaciones para su fabricación, de forma que cuanto menor sea el nivel de tolerancia (esto es, mayor exigencia), tanto mejor será la calidad de proyecto, y viceversa. No obstante, el proyecto debe ser lo más económico y sencillo posible, cumpliendo las expectativas del consumidor.

Una vez fabricado el artículo, el grado en el que éste concuerda con las exigencias del proyecto original, es decir, con el Control de Calidad tanto en los materiales utilizados como en los productos que salen de la fábrica

y se almacenan en las instalaciones de la empresa, se denomina *calidad de concordancia con el proyecto*.

El funcionamiento final de un producto depende tanto de la calidad de proyecto como de la calidad de concordancia. Aun cuando un proyecto sea inmejorable, si el control de concordancia ha sido inadecuado, por ejemplo, porque no se han utilizado los materiales consignados en la planificación original, es posible que el artículo no funcione correctamente. Por otra parte, aunque el control de concordancia fuese excepcional, si el artículo está inicialmente mal proyectado, no se conseguirá un funcionamiento satisfactorio.

En consecuencia, es necesario establecer un sistema que informe sobre la calidad de todo el proceso, para así poder tomar decisiones que optimicen la calidad del producto.

2.4 Causas Naturales y Atribuibles

En un proceso de fabricación las causas de la variación de la calidad son, esencialmente, de dos tipos:

- *Causas naturales o fortuitas*: Todo proceso de fabricación presenta un cierto grado de variabilidad inherente o natural que algunos autores denominan *ruido de fondo*. Esta variabilidad natural es el efecto acumulado de muchas pequeñas causas, generalmente incontrolables, debidas al azar: son las *causas naturales*.
- *Causas atribuibles*: Son aquellas que no forman parte de la variabilidad natural del proceso, como pueden ser un ajuste inapropiado de las máquinas, errores del trabajador, defectos en las materias primas, etc. En otras palabras, son causas externas identificables que se pueden descubrir y evitar en el futuro. Generalmente, la variabilidad provocada por las causas atribuibles es mayor que la provocada por las causas naturales.

Cuando el ruido de fondo de un proceso es relativamente pequeño, suele considerarse correcto el funcionamiento del proceso. Por tanto, un proceso cuya variabilidad en el funcionamiento sólo está influenciada por causas naturales se considera *bajo control estadístico*. Por el contrario, un proceso que funciona en presencia de causas atribuibles se considera *fuera de control*.

Así pues, uno de los principales *objetivos* del control estadístico de procesos es detectar rápidamente las causas atribuibles, a fin de que se puedan analizar y corregir apropiadamente antes de que aparezcan muchos elementos que no cumplan las especificaciones. Para ello una herramienta básica es la construcción de gráficos de control que establecen el recorrido de las variaciones que no es probable que se produzcan por azar; por lo que cualquier variación que se produzca dentro de este recorrido se debe a una causa externa que hay que identificar.

2.5 Funcionamiento de un Gráfico de Control

2.5.1 Principios Básicos

Un *gráfico de control* es un dibujo para determinar si el modelo de probabilidad subyacente en el proceso es estable o cambia a lo largo del tiempo. En él se representan las realizaciones muestrales de una característica de calidad en función de la posición o del tiempo. El gráfico consta de tres líneas horizontales, paralelas y, generalmente¹, equidistantes:

- la *línea central (LC)*, que representa el valor medio de la característica de calidad, correspondiente al estado bajo control, es decir, únicamente en presencia de causas naturales;
- la *línea superior de control (LSC)*, situada encima de la línea central; y
- la *línea inferior de control (LIC)*, por debajo de la línea central.

Estos límites de control se construyen de forma que, si el proceso está bajo control, la mayoría de los puntos muestrales se encuentre entre ellos. En ese caso, no es necesario tomar ninguna medida, sino que dejamos que el proceso actúe sólo. Sin embargo, si más de $m\alpha$ puntos se encuentran fuera de los límites del gráfico (siendo m el número de puntos considerados y α la probabilidad de que uno de ellos salga fuera de los límites cuando

¹La construcción general de los límites de control propuesta por Shewart establece límites de control equidistantes. Sin embargo, estudiaremos también algunos gráficos de control en los que no se sigue la filosofía de este autor y los límites no equidistan de la línea central.

el proceso está bajo control), se interpreta como una evidencia de que el proceso está fuera de control y es necesario encontrar y eliminar las causas que producen este comportamiento. Obsérvese que se permite que, como máximo, $m\alpha$ puntos caigan fuera de los límites de control por azar. En la mayoría de los casos, $m\alpha$ será menor que 1, de modo que bastará con que un punto dé una condición fuera de control para afirmar que el proceso no funciona correctamente.

Los puntos muestrales representados en el gráfico de control suelen unirse mediante segmentos para permitir una visualización más clara de la evolución del proceso a lo largo del tiempo.

Por otra parte, si todos los puntos se hallan dentro de los límites de control, pero se comportan de forma no aleatoria, nos indica que el proceso está fuera de control.

Esencialmente, un gráfico de control es un contraste de hipótesis donde:

$$\left. \begin{array}{l} H_0 : \text{Proceso bajo control estadístico} \\ H_1 : \text{Proceso fuera de control estadístico} \end{array} \right\}$$

Así, la regla de decisión consiste en rechazar la hipótesis nula si un punto muestral se encuentra fuera de los límites; mientras que si dicho punto está comprendido entre los límites no tenemos evidencia suficiente para rechazar dicha hipótesis.

Al igual que en todo contraste de hipótesis, pueden cometerse dos tipos de errores:

- *Error de tipo I*, que consiste en concluir que el proceso está fuera de control estadístico cuando en realidad está bajo control.
- *Error de tipo II*, cuando se concluye que el proceso está bajo control estadístico cuando ciertamente no es así.

Recuérdese que α representa la probabilidad de cometer error de tipo I, mientras que β representa la probabilidad de cometer error de tipo II. Así pues, en Control de Calidad – específicamente en Muestreo para la Aceptación – α es la probabilidad de rechazar un lote de calidad aceptable, mientras que β representa la probabilidad de aceptar un lote con una proporción alta de artículos defectuosos.

Con objeto de ilustrar la construcción de los gráficos de control (posteriormente estudiaremos cómo se calculan los límites) se proponen los siguientes ejemplos.

Nº de botella	X (contenido, en l)	Nº de botella	X (contenido, en l)
1	0.94	11	1.02
2	0.92	12	0.94
3	0.96	13	1.06
4	1.02	14	1.00
5	0.95	15	0.98
6	1.05	16	1.03
7	0.99	17	0.94
8	1.00	18	0.92
9	1.00	19	1.01
10	0.99	20	1.04

Tabla 2.1: Datos sobre el contenido de agua de una muestra de 30 botellas

Ejemplo 23 *Una fábrica embotelladora de agua desea que sus botellas tengan un contenido de 1l. Se selecciona al azar una muestra de 20 botellas consecutivas y se mide la cantidad de agua que contienen. Los resultados obtenidos se recogen en la Tabla 2.1.*

La Figura 2.1 muestra el gráfico de control asociado a este ejemplo concreto. Se observa que todos los valores muestrales se encuentran dentro de las bandas de control y en torno a la línea central ($\bar{x} = 0.988$), en consecuencia, puesto que no se tiene ninguna razón que obligue a actuar de otro modo, es preferible dejar que el proceso de embotellado funcione sólo.

Ejemplo 24 *En una cadena de producción, un operario hace agujeros en una pieza de metal con un taladro. Las exigencias del proyecto en cuanto a la profundidad del orificio son de 2 ± 0.01 cm. Para controlar si la operación de taladrado es estable a lo largo del tiempo se miden los orificios realizados a 30 piezas, obteniéndose los resultados de la Tabla 2.2.*

Se construye el gráfico de control para la característica de calidad X : “Profundidad del orificio (en cm)” (véase Figura 2.2). En este ejemplo, la muestra número 26 correspondiente a la profundidad del orificio de la pieza que ocupa dicho lugar, escapa de los límites de variación natural del proceso de taladrado. Por tanto, el proceso está fuera de control.

Nótese que los límites de control no coinciden, en general, con los límites de especificación. Los primeros están asociados a la variabilidad natural inherente al proceso, mientras que los segundos no están determinados por

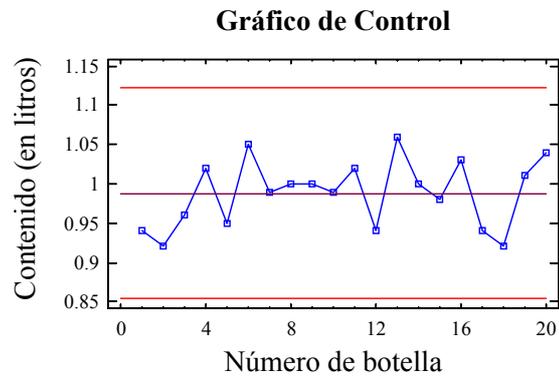


Figura 2.1: Gráfico de control para el contenido de agua en botellas

Nº de pieza	Profundidad (en cm)	Nº de pieza	Profundidad (en cm)
1	1.910	16	2.021
2	1.928	17	2.032
3	1.920	18	1.975
4	1.980	19	1.954
5	1.990	20	1.934
6	2.051	21	2.000
7	2.068	21	1.991
8	2.017	23	1.909
9	2.011	24	1.964
10	2.016	25	1.975
11	2.045	26	2.120
12	1.974	27	2.040
13	1.892	28	2.032
14	1.967	29	1.935
15	1.932	30	1.989

Tabla 2.2: Datos sobre la profundidad del orificio realizado por un taladro

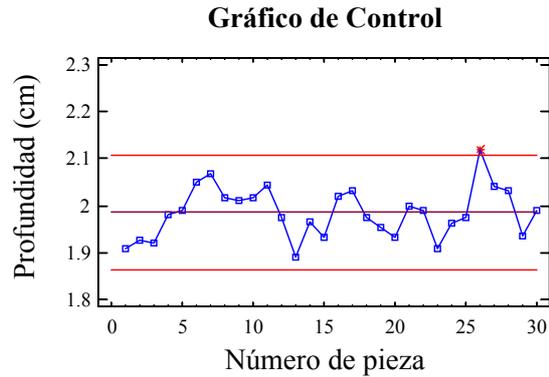


Figura 2.2: Gráfico de control para la profundidad del orificio de una pieza de metal

el proceso, sino que vienen impuestos desde fuera, por ejemplo, por la Administración o la Dirección de la Empresa. En este caso concreto, los límites de especificación son más restrictivos, pues están dentro de los de control.

2.5.2 Patrones de Comportamiento

Como se ha comentado en la sección anterior, aunque todos los puntos muestrales estén comprendidos entre los límites superior e inferior de control, si la disposición de estos puntos en el gráfico no es aleatoria sino que siguen un patrón sistemático, se dice que el proceso está fuera de control. En la Figura 2.3 se resumen algunos de los patrones de comportamiento no aleatorios que pueden presentarse en un gráfico de control.

2.5.3 Diseño de un Gráfico de Control

Son tres los aspectos más importantes a la hora de diseñar un gráfico de control y que se desarrollan con detalle en cada una de las subsecciones siguientes: seleccionar los límites de control apropiados para construir el gráfico de control, establecer el tamaño que han de tener las muestras elegidas y determinar con qué frecuencia deben extraerse.

TIPO DE PATRÓN	SIGNIFICADO	POSIBLES CAUSAS
CÍCLICO	El gráfico presenta ciclos (el comportamiento se repite).	<ul style="list-style-type: none"> • Cambios ambientales cíclicos (temperatura, fatiga del operario, etc.) • Fluctuación de voltaje o presión • Horarios de mantenimiento o desgaste de las herramientas
MEZCLADO	Los puntos tienden a quedar cerca o un poco fuera de los límites de control, con pocos puntos cercanos a la línea central.	“Sobrecontrol” cuando los operarios ajustan el proceso con demasiada frecuencia
CAMBIOS EN EL NIVEL	Se produce un cambio en el nivel del proceso.	<ul style="list-style-type: none"> • Introducción de nuevos trabajadores, métodos, materias primas, etc. • Cambios en la destreza, atención o motivación de los operarios
TENDENCIA	Un desplazamiento continuo en cierta dirección.	<ul style="list-style-type: none"> • Desgaste o deterioro gradual de una herramienta o componente del proceso • Cansancio del operador o presencia del supervisor • Influencias estacionales
ESTRATIFICACIÓN	Tendencia de los puntos a estar agrupados artificialmente alrededor de la línea central.	<ul style="list-style-type: none"> • Cálculo incorrecto de los límites de control • El proceso de muestreo recoge una o más unidades de varias distribuciones distintas subyacentes

Figura 2.3: Patrones de comportamiento no aleatorio

Selección de los Límites de Control

Para especificar los límites de control se ha de tener en cuenta lo siguiente:

- Si dichos límites se alejan de la línea central, se reduce el riesgo de cometer error de tipo I, puesto que será menos probable que un punto caiga fuera de los límites cuando el proceso está bajo control estadístico; sin embargo, el riesgo de cometer error de tipo II se incrementa, ya que los puntos que deberían estar fuera de los límites de control (nótese que el proceso está fuera de control) ahora quedarán dentro, lo que conducirá a la aceptación de control estadístico.
- Por el contrario, si estos límites se acercan a la línea central, se reduce el riesgo de cometer error de tipo II, pero aumenta la probabilidad de cometer error de tipo I.

En consecuencia, hay que llegar a un equilibrio entre ambos, análogo al que se emplea en la construcción de los intervalos de confianza y contrastes de hipótesis, esto es, fijar α en un valor pequeño y minimizar β .

El objetivo es probar si un proceso es estable o si, por el contrario, cambia a lo largo del tiempo. Por consiguiente, si la distribución de la característica de calidad de interés es, aproximadamente, Normal con parámetros μ_0 y σ , se desea probar la hipótesis de que la media de dicha distribución es estable, esto es

$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{array} \right\} \quad (2.1)$$

donde μ_0 es la media de la distribución. Si la desviación típica σ de la característica de calidad es conocida, el estadístico de contraste asociado al test (2.1) se construye a partir de una muestra aleatoria simple X_1, \dots, X_m y viene dado por

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{m}}} \rightsquigarrow \mathcal{N}(0, 1),$$

de modo que aquellos valores del estadístico experimental comprendidos entre $z_{\frac{\alpha}{2}}$ y $z_{1-\frac{\alpha}{2}}$, es decir, entre los cuantiles de una $\mathcal{N}(0, 1)$ de órdenes $\frac{\alpha}{2}$ y $1 - \frac{\alpha}{2}$, respectivamente, conducen a la aceptación de la hipótesis nula y, por

tanto, de estabilidad de la media (o control estadístico). En resumen, se acepta la hipótesis nula siempre y cuando la media muestral esté contenida en el intervalo

$$\left[\mu_0 - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{m}}, \mu_0 + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{m}} \right]$$

En cualquier caso, se acepta la hipótesis nula si

$$\bar{x} \in \left[\mu_0 - k \frac{\sigma}{\sqrt{m}}, \mu_0 + k \frac{\sigma}{\sqrt{m}} \right]$$

procediendo a la estimación de los parámetros si alguno de ellos es desconocido. Es claro que la constante k está estrechamente relacionada con α

$$k = z_{1-\frac{\alpha}{2}}.$$

Ésta es la idea intuitiva de cómo se construyen los gráficos de control.

Así pues, la construcción más usual de límites de control, debida a Shewart, es la siguiente. Sea W un estadístico asociado a la característica de calidad X objeto de estudio, con media μ_W y desviación típica σ_W . Entonces, los límites de control se calculan como:

$$\begin{aligned} LSC &= \mu_W + k\sigma_W \\ LC &= \mu_W \\ LIC &= \mu_W - k\sigma_W, \end{aligned}$$

donde k es una constante que determina la distancia de las líneas de control a la línea central. Nótese que, fijado un valor k , puede obtenerse la probabilidad α de cometer error de tipo I mediante la relación²:

$$1 - \alpha = P[-k < Z < k] \Leftrightarrow k = z_{1-\frac{\alpha}{2}}.$$

De esta forma, los límites de control pueden calcularse fijando k o bien, para un α dado, obtener el valor k asociado y, a partir de él, construir los límites del gráfico. Éstos últimos se denominan *límites probabilísticos de α* . En E.E.U.U. se suelen determinar límites de control tomando $k = 3$, mientras que en algunas zonas de Europa occidental se opta por el cálculo de límites probabilísticos para $\alpha = 0.002$.

²Para ello se supone que $W \rightsquigarrow \mathcal{N}(\mu_W, \sigma_W)$, de lo contrario, $1 - \alpha = P[-k < W < k]$.

Ejemplo 25 Retómese la operación de taladrado del ejemplo anterior. El objetivo es establecer un gráfico de control para determinar si dicho proceso funciona bajo control estadístico. Para ello se toma como estadístico asociado a la característica de calidad, la propia característica, cuya media y varianza son desconocidas. No obstante, es posible estimar dichos valores a partir de la muestra extraída. Para ello se utilizan los estimadores puntuales habituales:

$$\hat{\mu} = \bar{x} = \frac{59.572}{30} = 1.9857$$

$$\hat{\sigma} = S = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2} = 0.0533,$$

de donde los límites de control son:

$k = 3$)

$$LIC = \bar{x} - 3 \cdot S = 1.9857 - 3 \cdot 0.0533 = 1.8258$$

$$LC = \bar{x} = 1.9857$$

$$LSC = \bar{x} + 3 \cdot S = 1.9857 + 3 \cdot 0.0533 = 2.1456.$$

Para determinar el valor de α :

$$3 = z_{1-\frac{\alpha}{2}} \Rightarrow 1 - \frac{\alpha}{2} = 0.9987 \Rightarrow \alpha = 0.0027.$$

En consecuencia, el proceso está fuera de control si más de $m\alpha = 30 \cdot 0.0027 = 0.081 \cong 0$ muestras están fuera de los límites obtenidos. Obsérvese que ningún valor sobrepasa dichos límites, aunque la muestra número 26 se sitúa cerca del límite superior de control, lo que lleva a la conclusión de que el proceso está bajo control estadístico.

$k = 2$)

$$LIC = \bar{x} - 2 \cdot S = 1.8791$$

$$LC = \bar{x} = 1.9857$$

$$LSC = \bar{x} + 2 \cdot S = 2.0923.$$

Y, ahora, la probabilidad de cometer error de tipo I es:

$$2 = z_{1-\frac{\alpha}{2}} \Rightarrow 1 - \frac{\alpha}{2} = 0.9772 \Rightarrow \alpha = 0.0455.$$

En este caso, el número máximo de muestras que, por azar, pueden caer fuera de los límites de control es $m\alpha = 30 \cdot 0.0455 = 1.245 \cong 1$. En este caso, la observación número 26 se sitúa por encima del límite superior de control ($x_{26} = 2.120$), lo cual es predecible, pues al disminuir k la amplitud de los límites de control disminuye. No obstante, no se puede estar razonablemente seguro de que el proceso está fuera de control, pues es posible que la causa de que dicha muestra dé una señal sea natural y no atribuible.

Si se desea determinar límites de control probabilísticos para $\alpha = 0.002$:

$$k = z_{1-\frac{0.002}{2}} = z_{0.999} = 3.09,$$

luego

$$LIC = \bar{x} - 3.09 \cdot S = 1.8210$$

$$LC = \bar{x} = 1.9857$$

$$LSC = \bar{x} + 3.09 \cdot S = 2.1504.$$

Ha aumentado ligeramente el valor de k , con lo que se incrementa la amplitud de los límites de control y, por tanto, el proceso se mantiene bajo control estadístico.

Tamaño Muestral y Frecuencia de Muestreo

El tamaño y la frecuencia de muestreo dependerán, en general, del cambio en el proceso que se quiera detectar. Las muestras grandes permiten detectar más fácilmente cambios pequeños, mientras que las muestras pequeñas se emplean para la detección de cambios grandes. No obstante, lo deseable sería extraer muestras pequeñas - siempre y cuando éstas fuesen representativas - a intervalos largos de tiempo, es decir, con menor frecuencia, o bien muestras grandes con mayor asiduidad. El inconveniente es que esto supone un elevado costo económico y temporal. Por tanto, se opta por extraer muestras de mayor tamaño con menor frecuencia o bien, muestras pequeñas poco espaciadas en el tiempo, distribuyendo de esta forma el esfuerzo y el coste de muestreo. Posteriormente, se estudia un método para el cálculo del tamaño muestral a partir de la curva OC.

Subgrupos Racionales

El concepto de subgrupos racionales se refiere a que, si el proceso funciona en presencia de causas atribuibles, las diferencias que puedan existir entre las muestras y dentro de ellas, sean máxima y mínima, respectivamente. Para establecer subgrupos racionales se emplean, básicamente, dos enfoques:

1. Si el objetivo es detectar cambios en el proceso, se toman muestras que consten de artículos que se produjeron simultáneamente o con la menor diferencia posible.
2. Si se pretende tomar decisiones acerca de la aceptación de todos los artículos que se han producido desde la última muestra, lo más conveniente es que cada muestra sea una muestra aleatoria de todos los artículos fabricados en el intervalo de muestreo. Éstos serán representativos de todos los producidos desde que se formó el último subgrupo racional.

Asimismo, se combinan ambos métodos cuando alguno de ellos no proporciona los resultados deseados.

Con esta sección se pone de manifiesto la importancia de establecer muestras, pues son éstas las que permiten la extracción de conclusiones apropiadas acerca del funcionamiento de un proceso.

Límites de Tolerancia Natural

Los límites de tolerancia natural de un proceso son los límites determinados por la variabilidad natural medida por la desviación típica del proceso, denominada σ , es decir

$$\begin{aligned}LSC &= \mu + k\sigma \\LC &= \mu \\LIC &= \mu - k\sigma.\end{aligned}$$

Estos límites no tienen relación con los *límites de especificación* del proceso ya que estos últimos vienen dados externamente por la Administración, el consumidor, el fabricante, etc. Estabilidad del Proceso

Para analizar la estabilidad de un proceso de fabricación y así asegurar la calidad de los productos fabricados, se establecen, generalmente, dos períodos: el *período base* o *período de prueba* y el *período de vigilancia*. En el período base se establecen límites de control de prueba a partir de un conjunto de observaciones. Si durante dicho período el proceso está bajo control, es decir, es estable, podrán emplearse con toda confianza los límites de control anteriormente obtenidos para continuar controlando el proceso durante el período de vigilancia.

La inestabilidad puede deberse, fundamentalmente, a dos causas:

- Que sea inherente al proceso (equipo en mal estado, baja calidad de los materiales, etc.)
- Debido a la novedad del producto o del proceso (los trabajadores encargados no están familiarizados).

¿Cómo se establecen entonces los límites de control? Son dos las posibles opciones:

1. Si el tiempo y el costo lo permiten, se puede iniciar un nuevo período base. Probablemente el proceso ya se habrá estabilizado.
2. Controlar si el proceso es estable al eliminar las observaciones que se encuentran fuera de los límites de control. Para ello se siguen los siguientes pasos:

Paso 1) Iniciar un período base de, al menos, 20 muestras³.

Paso 2) Calcular los límites de control del gráfico considerado a partir de las muestras anteriores; estos límites se denominan *límites de control de prueba*.

Paso 3) Revisar el gráfico, de forma que si todas las observaciones caen dentro de los límites de control, vamos al Paso 4; de lo contrario, y bajo el supuesto de presencia de causas atribuibles, habrá que eliminar las muestras implicadas del período base y volver de nuevo al Paso 2.

Paso 4) Extender los límites de control al período de vigilancia.

³Bertrand (1990) recomienda la utilización de un período base con, al menos, 20 muestras.

En resumen, se tiene el diagrama de la Figura 2.4.

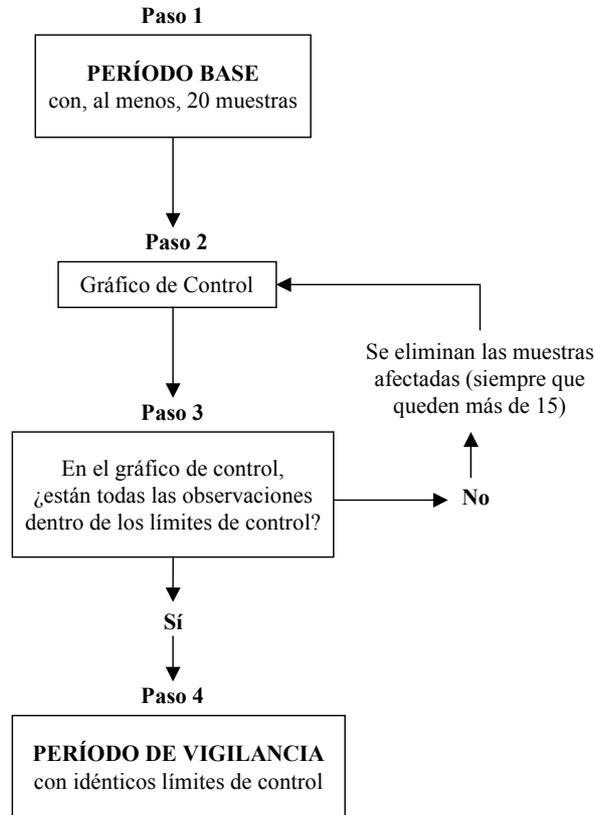


Figura 2.4: Pasos para estabilizar un proceso

Nota 26 *Es claro que no se pueden eliminar muestras de forma indiscriminada. Si quedasen menos de 15, se parte de un nuevo conjunto de muestras o bien se prescinde del período de prueba.*

2.6 Límites de aviso

Los límites de control de un gráfico convencional dan una señal indicando que el proceso puede estar fuera de control. En una distribución Normal

con los límites de control para $k = 3$, la probabilidad de una señal falsa es de sólo 0.0027 (error tipo I)⁴. De esta manera, se puede estar razonablemente seguro de que el proceso está fuera de control cuando se produzca la señal. No se estará seguro si no se produce la señal; incluso aunque el proceso haya pasado a estar fuera de control, es posible que la señal no se produzca de inmediato. Para intentar evitar estos retrasos indebidos, algunos expertos recomiendan el uso de límites de aviso o advertencia. Éstos se encuentran situados dentro de los límites usuales de control, de modo que si los límites iniciales se construyen para $k = 3$ o probabilísticos de 0.002, los límites de aviso se calculan con $k = 2$ o de 0.05, respectivamente. Así, si se observa que un punto está fuera de los límites de aviso, pero sigue situado dentro de los de control, hay que sospechar que el proceso no funciona de forma apropiada y, en consecuencia, se ha de vigilar el proceso con mayor atención. Una alternativa consiste en incrementar la frecuencia de muestreo y emplear estos nuevos datos junto con los puntos sospechosos para investigar si el proceso está bajo control.

Si varias muestras sucesivas están dentro de los límites de aviso no hay porqué preocuparse. La norma siempre ha de ser: *Deje al proceso sólo a no ser que haya una razón clara para intervenir.*

La probabilidad de que dos muestras sucesivas estén fuera de los límites de aviso es muy reducida. Sean x_1 y x_2 dos valores observados de la muestra que están fuera de los límites de aviso de 2σ . Entonces, aplicando que ambos son independientes y suponiendo normalidad, se tiene:

$$\begin{aligned} P[x_1 \text{ esté fuera de los límites}] &= 1 - P[-2 < Z < 2] = 0.0456 \\ P[x_2 \text{ esté fuera de los límites}] &= 0.0455 \\ P[x_1 \text{ y } x_2 \text{ estén fuera de los límites}] &= P[x_1 \text{ esté fuera}] \cdot P[x_2 \text{ esté fuera}] \\ &= 0.0455^2 = 0.00207. \end{aligned}$$

Esta probabilidad es incluso más pequeña que la de recibir una señal en un gráfico de \bar{X} con límites de 3σ (0.0027). Por tanto, si el proceso da un aviso y la siguiente muestra lo repite, el proceso funciona bajo la presencia de causas atribuibles.

Los límites de aviso permiten que el gráfico de control sea más sensible a los cambios en los parámetros del proceso, el inconveniente es que su

⁴ $\alpha = P[\text{Señal falsa}] = 1 - P[-3 < Z < 3] = 0.0027.$

interpretación no es demasiado precisa y, en ocasiones, pueden llevar a equívocos.

2.7 Curva Característica Operativa

La probabilidad de aceptación de una muestra no es más que la probabilidad de asumir que el proceso está bajo control. Esta probabilidad dependerá, en cada caso, de si realmente el proceso está bajo control o no. Así pues, dado el contraste de hipótesis

$$\left. \begin{array}{l} H_0 : \text{Proceso bajo control} \\ H_1 : \text{Proceso fuera de control} \end{array} \right\}$$

la probabilidad de aceptación viene dada por

$$P[\text{Aceptación}] = \begin{cases} 1 - \alpha & \text{bajo } H_0 \\ \beta & \text{bajo } H_1 \end{cases}$$

donde α y β denotan, como es habitual, las probabilidades de cometer error de tipo I y II, respectivamente. Esta probabilidad de aceptación variará de acuerdo con el cambio que se produzca en el proceso de producción, por ejemplo, si el desajuste se produce en la media, será una función de la diferencia existente entre la verdadera media del proceso y la media bajo H_0 , de modo que cuanto mayor sea dicha diferencia, tanto menor será la probabilidad de aceptación y viceversa.

En consecuencia, la *Curva Característica Operativa*, también denominada *Curva Característica de Operación* o *Curva OC*, asociada a un gráfico de control, es una representación gráfica de la probabilidad de cometer error de tipo II, β . Esta función dependerá, en general, del tamaño muestral considerado y del cambio producido en el proceso, esto es

$$\beta = f(n, \alpha, d)$$

donde d representa dicha alteración.⁵

Un fácil procedimiento para su cálculo se describe a continuación.

⁵Notemos que la probabilidad de cometer error de tipo II está estrechamente ligada con la probabilidad de cometer error de tipo I.

Sean θ_1 y θ_0 el valor real del parámetro de interés y bajo la hipótesis nula, respectivamente. Entonces, el cambio producido viene dado por:

$$d = \theta_1 - \theta_0.$$

Por tanto, los pasos a seguir son:

P1) Se establecen límites superior e inferior del gráfico de control correspondiente, dado un valor k (generalmente $k = 3$).

P2) Se obtiene la desviación típica σ_W a partir de los límites como

$$\sigma_W = \frac{LSC - LIC}{2k}.$$

P3) Se selecciona la diferencia d . Usualmente los valores de d se toman múltiplos de σ_W con objeto de facilitar los cálculos⁶.

P4) Se calcula la probabilidad de aceptación de una muestra para cada diferencia d .

Ejemplo 27 *Los límites de control establecidos para vigilar la estabilidad de la longitud media de cable eléctrico fueron 25 ∓ 5 . Para ello, se extrajeron varias muestras de tamaño 6 y se calculó la media muestral para cada una de ellas. Por tanto, el estadístico a partir del cuál se construyeron los límites de control para la media fue $W = \bar{X}$.*

El uso del gráfico de control equivale al contraste de hipótesis siguiente:

$$\left. \begin{array}{l} H_0 : \mu = 25 \\ H_1 : \mu \neq 25 \end{array} \right\}$$

Puede prescindirse del Paso 1 puesto que el ejemplo ya proporciona los límites de control. En el Paso 2, se obtiene la desviación típica $\sigma_{\bar{X}}$:

$$\sigma_{\bar{X}} = \frac{30 - 20}{6} = \frac{5}{3}.$$

Así pues, las diferencias d serán, por ejemplo, $0, \frac{5}{3}, \frac{10}{3}, 5, \frac{20}{3}, \frac{25}{3}, 10$.

Seguidamente se calcula la probabilidad de aceptación para cada valor de d :

⁶Estas diferencias pueden ser negativas si d es negativo, sin embargo no suelen considerarse porque la curva OC es simétrica con respecto a $d = 0$.

$$d = 0 \Leftrightarrow \mu_1 = \mu_0 = 25 :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \rightsquigarrow \mathcal{N} \left(25, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - 25}{\frac{5}{3}} < Z < \frac{30 - 25}{\frac{5}{3}} \right] = P[-3 < Z < 3] = 0.9974. \end{aligned}$$

$$d = \frac{5}{3} \Leftrightarrow \mu_1 = \mu_0 + d = \frac{80}{3} :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \rightsquigarrow \mathcal{N} \left(\frac{80}{3}, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - \frac{80}{3}}{\frac{5}{3}} < Z < \frac{30 - \frac{80}{3}}{\frac{5}{3}} \right] = P[-4 < Z < 2] = 0.9772. \end{aligned}$$

$$d = \frac{10}{3} \Leftrightarrow \mu_1 = \mu_0 + d = \frac{85}{3} :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \rightsquigarrow \mathcal{N} \left(\frac{85}{3}, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - \frac{85}{3}}{\frac{5}{3}} < Z < \frac{30 - \frac{85}{3}}{\frac{5}{3}} \right] = P[-5 < Z < 1] = 0.8413. \end{aligned}$$

$$d = 5 \Leftrightarrow \mu_1 = \mu_0 + d = 30 :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \rightsquigarrow \mathcal{N} \left(30, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - 30}{\frac{5}{3}} < Z < \frac{30 - 30}{\frac{5}{3}} \right] = P[-6 < Z < 0] = 0.5000. \end{aligned}$$

$$d = \frac{20}{3} \Leftrightarrow \mu_1 = \mu_0 + d = \frac{95}{3} :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \rightsquigarrow \mathcal{N} \left(\frac{95}{3}, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - \frac{95}{3}}{\frac{5}{3}} < Z < \frac{30 - \frac{95}{3}}{\frac{5}{3}} \right] = P[-7 < Z < -1] = 0.1587. \end{aligned}$$

$$d = \frac{25}{3} \Leftrightarrow \mu_1 = \mu_0 + d = \frac{100}{3} :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \sim \mathcal{N} \left(\frac{100}{3}, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - \frac{100}{3}}{\frac{5}{3}} < Z < \frac{30 - \frac{100}{3}}{\frac{5}{3}} \right] = P[-8 < Z < -2] = 0.0228. \end{aligned}$$

$$d = 10 \Leftrightarrow \mu_1 = \mu_0 + d = 35 :$$

$$\begin{aligned} P[\text{Aceptacin}] &= P \left[LIC < \bar{X} < LSC \mid \bar{X} \sim \mathcal{N} \left(35, \frac{5}{3} \right) \right] = \\ &= P \left[\frac{20 - 35}{\frac{5}{3}} < Z < \frac{30 - 35}{\frac{5}{3}} \right] = P[-9 < Z < -3] = 0.0014. \end{aligned}$$

La Figura 2.5 muestra la representación gráfica de la curva OC.

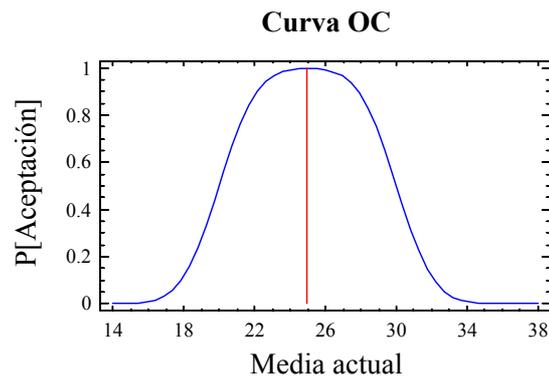


Figura 2.5: Curva OC para la longitud de cable eléctrico

2.7.1 La Curva OC y el Tamaño Muestral

Obsérvese que la relación entre d y σ_W es crucial para estudiar el comportamiento de la curva característica operativa, ya que:

- si $\frac{d}{\sigma_W} \rightarrow \infty \Rightarrow P[\text{Aceptación}] \rightarrow 0$.
- si $\frac{d}{\sigma_W} \rightarrow 0 \Rightarrow P[\text{Aceptación}] \rightarrow 1 - \alpha$.

Por otra parte, centrándose en el estudio de la media del proceso y extrayendo muestras de tamaño n , el estadístico W será \bar{X} . Entonces, puesto que $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, al aumentar el tamaño muestral n , disminuye la desviación típica $\sigma_{\bar{X}}$, con lo que aumenta el cociente $\frac{d}{\sigma_{\bar{X}}}$ y, por tanto, disminuye la probabilidad de aceptación. En consecuencia, al aumentar el tamaño de las muestras, la curva OC, en este caso, resulta más pronunciada; por el contrario, si el tamaño de las muestras se hace pequeño, la curva OC tenderá a ser una línea horizontal. Esta relación entre la curva OC y el tamaño de la muestra posibilita la obtención del tamaño mínimo que debe tener una muestra, fijada la probabilidad de detección de un cierto cambio en la media del proceso.

Ejemplo 28 *Una planta diseña un nuevo componente para máquinas agrícolas y se dispone a su fabricación. El director de dicha planta desea saber si los estándares de calidad son óptimos y estables. Para ello se mide una cierta característica de calidad del componente cuya media se espera que sea igual a 10. Si se sabe además que la desviación típica del proceso es 0.5 ¿cuál será el tamaño muestral mínimo necesario para detectar un cambio de 1 unidad hacia arriba en la media, con probabilidad 0.95, justo en la primera muestra después del cambio?*

Como se quiere detectar un cambio de 1 unidad hacia arriba en la media, se lleva a cabo el siguiente contraste:

$$\left. \begin{array}{l} H_0 : \mu = 10 \\ H_1 : \mu = 11 \end{array} \right\}$$

Así, $W = \bar{X}$ y

$$\begin{aligned} P[\text{Detección}] &= P\left[\bar{X} > LSC \mid \bar{X} \sim \mathcal{N}\left(11, \frac{0.5}{\sqrt{n}}\right)\right] = \\ &= P\left[\bar{X} > 10 + 3\frac{0.5}{\sqrt{n}} \mid \bar{X} \sim \mathcal{N}\left(11, \frac{0.5}{\sqrt{n}}\right)\right] = \\ &= P\left[Z > \frac{10 + 3\frac{0.5}{\sqrt{n}} - 11}{\frac{0.5}{\sqrt{n}}}\right] = P\left[Z > 3 - \frac{\sqrt{n}}{0.5}\right]. \end{aligned}$$

Entonces

$$\begin{aligned} P[\text{Aceptación}] &= 1 - P[\text{Detección}] = 1 - 0.95 = 0.05 \\ P[\text{Aceptación}] &= P\left[Z \leq 3 - \frac{\sqrt{n}}{0.5}\right] = 0.05 \end{aligned}$$

En consecuencia, buscando el cuantil de orden 0.05 de una $\mathcal{N}(0, 1)$, esto es

$$z_{0.05} = -1.645,$$

se tiene el tamaño pedido

$$3 - \frac{\sqrt{n}}{0.5} = z_{0.05} = -1.645 \Leftrightarrow n = [0.5(3 + 1.645)]^2 = 5.3940.$$

Por tanto, se precisan muestras que contengan un número mínimo de 6 elementos.

2.7.2 La Curva OC y la Longitud Media de Racha

Una *racha* es el conjunto de muestras que es necesario inspeccionar para que se detecte una señal de que el proceso está fuera de control en el gráfico correspondiente. Entonces, la *longitud de racha* se define como el número de muestras que componen dicha racha. Así, por ejemplo, si en un gráfico de control se produce una señal en la 5ª muestra observada, la racha está formada por las muestras 1, 2, 3, 4 y 5 y la longitud de racha es 5.

Considérese entonces la variable aleatoria J que representa la longitud de racha. Su función masa de probabilidad viene dada por

$$P[J = j] = P[Y = j - 1] = \theta(1 - \theta)^{j-1}, \quad j = 1, 2, 3, \dots$$

donde $Y \rightsquigarrow BN(1, \theta)$ con θ la probabilidad de detectar un cambio en el proceso en la primera muestra después del cambio. Obsérvese que la variable aleatoria Y representa el “número de muestras correctas antes de que se produzca la primera señal”.

Su función de distribución es, por tanto

$$\begin{aligned} F(j) &= P[J \leq j] = \sum_{k=1}^j P[J = k] = \sum_{k=1}^j \theta (1 - \theta)^{k-1} = \theta \sum_{k=1}^j (1 - \theta)^{k-1} = \\ &= \theta \frac{(1 - \theta)^j - (1 - \theta)^0}{(1 - \theta) - 1} = 1 - (1 - \theta)^j, \quad j = 1, 2, 3, \dots \end{aligned}$$

y su esperanza

$$E[J] = 1 + E[Y] = 1 + \frac{1 - \theta}{\theta} = \frac{1}{\theta},$$

que se conoce como *longitud media de racha* y se nota *ARL*.

Obsérvese que la curva OC y la longitud media de racha, ARL, están estrechamente relacionadas entre sí: la primera representa la probabilidad de aceptación de una muestra, esto es

$$P[\text{Aceptacin}] = 1 - P[\text{Deteccin}] = 1 - \theta,$$

mientras que la segunda es el inverso de la probabilidad de rechazo de la muestra (o de detección de la señal), $1/\theta$. En consecuencia, a medida que la probabilidad de aceptación de una muestra disminuye, la probabilidad de detección de una señal aumenta y, por consiguiente, la longitud media de racha disminuye hasta un valor mínimo igual a 1. Esto parece lógico, ya que habrá que inspeccionar menos muestras para detectar el cambio. Si, por el contrario, la probabilidad de aceptación de una muestra aumenta, esto implica que el número de muestras que habrá que extraer para detectar un cambio será mayor, con lo que la longitud media de racha aumentará. El resumen aparece en el siguiente cuadro:

<p>Si $OC = 1 - \theta \rightarrow 0 \Rightarrow \theta \rightarrow 1 \Rightarrow ARL = 1/\theta \rightarrow 1$ Si $OC = 1 - \theta \rightarrow 1 \Rightarrow \theta \rightarrow 0 \Rightarrow ARL = 1/\theta \rightarrow \infty$</p>
--

Existen tablas para calcular la longitud media de racha en un gráfico para la media. Aquí se incluyen dos de ellas, cuando la desviación típica del proceso σ es conocida (Tabla 2.3) y cuando ésta es desconocida pero se estima a partir de los recorridos muestrales (Tabla 2.4).

n	d/σ						
	0.1	0.2	0.5	1.0	1.5	2.0	3.0
3	314.8	223.9	54.7	9.1	2.9	1.5	1.0
4	303.9	192.1	42.9	6.3	2.0	1.2	1.0
5	289.3	175.6	33.4	4.5	1.6	1.1	1.0
6	277.8	159.0	26.3	3.4	1.3	1.0	1.0
7	267.5	147.5	21.3	2.8	1.2	1.0	1.0
8	258.9	133.0	17.7	2.3	1.1	1.0	1.0
9	251.3	122.0	15.0	2.0	1.1	1.0	1.0
10	244.0	111.1	12.8	1.8	1.0	1.0	1.0
15	203.9	76.4	6.9	1.2	1.0	1.0	1.0
20	175.6	56.6	4.5	1.1	1.0	1.0	1.0
25	155.5	41.1	3.2	1.0	1.0	1.0	1.0

Tabla 2.3: ARL (σ conocida)

n	d/\bar{R}					
	0.05	0.10	0.20	0.30	0.50	1.00
3	326.2	254.3	126.1	58.5	15.9	2.1
4	300.3	192.4	67.5	25.7	5.8	1.2
5	277.0	151.6	39.9	13.3	2.9	1.0
6	244.1	114.4	25.2	7.8	1.8	1.0
7	217.6	89.2	17.1	5.1	1.4	1.0
8	197.6	71.1	12.2	3.6	1.2	1.0
9	177.5	57.1	9.0	2.7	1.1	1.0
10	160.0	36.6	6.8	2.1	1.0	1.0
15	99.8	20.3	2.6	1.2	1.0	1.0
20	65.7	10.9	1.6	1.0	1.0	1.0
25	46.0	6.7	1.2	1.0	1.0	1.0

Tabla 2.4: ARL (σ desconocida, $\hat{\sigma} = \bar{R}/d_2$)

2.8 Ejercicios

- Los paquetes de galletas tradicionales vienen marcados con un contenido de peso neto de 500 g. Se tomó una muestra de 25 paquetes y se pesaron. Las medidas que se obtuvieron se recogen a continuación:

Nº de paquete	Peso (g)	Nº de paquete	Peso (g)
1	500	14	514
2	502	15	515
3	505	16	498
4	512	17	499
5	499	18	497
6	497	19	500
7	505	20	501
8	503	21	496
9	475	22	478
10	503	23	504
11	500	24	506
12	508	25	500
13	526		

Se pide:

- Defina la característica variable de calidad objeto de estudio.
 - Suponiendo que el peso de los paquetes se distribuye de forma Normal, estime la media y la varianza.
 - ¿A qué porcentaje de paquetes les faltarán más de 10 g de peso?
 - Determine los límites superior e inferior de control usando $k = 3$.
 - ¿Puede afirmarse que el proceso está bajo control estadístico?
- Una determinada característica de calidad se vigila mediante un gráfico de control diseñado de forma que la probabilidad de que se detecte una señal de fuera de control después del cambio hacia tal estado, es $1 - \beta$. Calcule:
 - La probabilidad de que se detecte que el proceso está fuera de control en la segunda muestra después del cambio.

- (b) La probabilidad de que se detecte que el proceso está fuera de control en la j -ésima muestra después del cambio.
3. Se construye el gráfico de control con límites de $3 - \text{sigmas}$ para una característica de calidad X a partir de 10 muestras extraídas al azar. Si se concluye que el proceso está fuera de control cuando una o más de las 10 muestras extraídas se sitúan fuera de los límites de control, calcule la probabilidad de cometer error de tipo I.
4. El director de una fábrica textil desea analizar si el proceso de fabricación de una nueva fibra sintética se mantiene bajo control. Para ello se realizaron las siguientes mediciones sobre la resistencia al desgaste de la fibra:

Medición	(psi, lb/plg ²)Resistencia	Medición	(psi, lb/plg ²)Resistencia
1	49.8	16	47.4
2	48.5	17	47.9
3	47.6	18	46.5
4	50.1	19	44.2
5	54.3	20	50.3
6	47.0	21	51.4
7	52.7	22	50.8
8	50.8	23	50.5
9	55.4	24	49.6
10	49.9	25	47.8
11	47.4	26	51.1
12	45.1	27	50.2
13	46.7	28	48.3
14	46.6	29	46.9
15	48.6	30	50.0

- (a) Estime la media y la desviación estándar de la resistencia de la fibra sintética.
- (b) Determine los límites superior e inferior de control utilizando $k = 2.5$. ¿Puede afirmarse que el proceso es estable?
- (c) Dibuje un gráfico de control simple. ¿Qué conclusiones se extraen del gráfico?

5. Se prueban respecto a su dureza 20 grados sucesivos de una aleación de acero. Los datos resultantes aparecen a continuación. Obtenga un diagrama de control para la dureza:
- usando $k = 3$.
 - usando límites probabilísticos con $\alpha = 0.001$.
 - Suponiendo causas atribuibles, establezca los límites de control para el período de vigilancia en (a) y (b).
 - ¿Cuáles son los parámetros estándares del proceso? Construya los límites de tolerancia natural para la dureza.

Grado	Dureza (codificada)
1	0.806
2	0.814
3	0.810
4	0.820
5	0.819
6	0.815
7	0.817
8	0.810
9	0.809
10	0.812
11	0.819
12	0.821
13	0.850
14	0.800
15	0.810
16	0.799
17	0.815
18	0.814
19	0.818
20	0.840

- Un gráfico de control usa límites de control de $\pm 2.5\sigma$ y límites de aviso de $\pm 2\sigma$. ¿Cuál es la probabilidad de que una observación genere un aviso pero no una señal?
- La longitud media de un regulador de glucosa para detectar la diabetes, producido en una operación estable, es 18 cm. Se supone que

- la desviación estándar es conocida y equivale a 0.1 cm. Se propone emplear un gráfico de control de \bar{X} para vigilar la longitud. ¿Cuál debería ser el tamaño de los subgrupos si ha de detectarse un descenso en la media de 0.2 cm con una probabilidad de 0.90, cuando se inspeccione el primer subgrupo después del cambio? ¿Y si esta probabilidad se incrementase hasta 0.95?
8. Se sabe que la desviación estándar de la fuerza de sujección de cierto tipo de clavos es de 0.04 *Newton*. Utilizando los gráficos de control de \bar{X} convencionales, con un tamaño de muestra de 10, realizándose muestreos cada hora, ¿cuánto tiempo se necesitará para detectar un cambio de 0.02 *Newton* en la media del proceso?
 9. Construya la curva característica operativa de un gráfico de control cuyos límites para $k = 3$ son $LIC = 230$ y $LSC = 242$.
 10. Se inicia la producción de un componente recientemente proyectado. Para vigilar su longitud se prepara un gráfico de control de 3σ cuyos límites son 15.5 y 24.5. ¿Cuál es la probabilidad de que se detecte un cambio de 2 cm en la media del proceso en la primera muestra después del cambio? ¿Cuál es la probabilidad de detectar dicho cambio entre las 5 primeras muestras?

Capítulo 3

Gráficos de Control de Atributos

3.1 Introducción

No todas las características de calidad pueden ser representadas de manera adecuada mediante números, pues hay casos en los que cada artículo o producto inspeccionado se debe clasificar como *conforme* o *disconforme* respecto a las especificaciones para la característica de calidad considerada (también suele hablarse de artículos *defectuosos* o *no defectuosos*, aunque la anterior terminología está más extendida). Estas características de calidad se denominan *atributos*. Como ejemplo de atributo se puede considerar que un tornillo es conforme si su longitud está entre 1.9 y 2.1 cm., mientras que en caso contrario será disconforme.

El gráfico construido para el control de este tipo de medida de calidad, se denomina *Gráfico de Control de Atributos*. Los cuatro gráficos que se utilizan habitualmente son:

- *Gráfico p* , que se refiere a la fracción de artículos disconformes o defectuosos producidos en un proceso de fabricación.
- *Gráfico np* , que se refiere al número de artículos disconformes o defectuosos producidos en un proceso de fabricación.
- *Gráfico c* , donde se mide el número de disconformidades o defectos observados.

- *Gráfico u* , que analiza el número de defectos por unidad inspeccionada.

3.2 Gráfico de Control para la Fracción de Disconformes: Gráfico p

Se considera una inspección de calidad de un artículo del siguiente modo: se examinan de manera simultánea varias características de calidad y si una, o más de una, de ellas no está de acuerdo con un estándar predefinido, se clasifica el artículo como disconforme.

Entonces, se define la fracción (o proporción) de artículos disconformes, o *fracción disconforme*, como el cociente del número de artículos disconformes en una población entre el número total de artículos que contiene dicha población. En general, la fracción de disconformes se suele representar mediante un número decimal (en tanto por uno), aunque a veces se representa el porcentaje de disconformes, que es el producto por 100 de la fracción disconforme (en tanto por ciento).

Nota 29 Aunque se suele trabajar con la fracción de artículos disconformes, se podría analizar también la fracción de artículos conformes, con lo que se tendría un gráfico de control para el rendimiento del proceso.

Los principios estadísticos en los que se basa el diagrama de control de la fracción o proporción de artículos disconformes proceden de la distribución Binomial, puesto que si el proceso de producción funciona de manera estable, la probabilidad de que cualquier artículo no esté conforme con las especificaciones es una constante (p), y los artículos producidos de manera sucesiva son independientes. Entonces, la disconformidad de cada artículo producido es una variable aleatoria de Bernoulli, con parámetro p ¹. Si se selecciona una muestra aleatoria de n artículos del producto estudiado, y si X es la variable aleatoria que representa el "número de artículos disconformes en dicha muestra", entonces X sigue una distribución Binomial de parámetros n y p :

$$X \sim B(n, p)$$

$$P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x} \quad x = 0, 1, \dots, n.$$

¹Se recuerda que la distribución de Bernoulli es la distribución Binomial (1, p)

La media de X es np y la varianza $np(1-p)$.

El objetivo del gráfico p es calcular y controlar la proporción o porcentaje de artículos defectuosos, es decir, p . Una estimación para p a partir de la muestra es la fracción disconforme muestral, que se define como el cociente del número de artículos disconformes X en la muestra, entre el número de artículos de la muestra (n), es decir:

$$\hat{p} = \frac{X}{n}.$$

Se puede obtener la distribución de la variable aleatoria \hat{p} a partir de la distribución Binomial de X , mediante la expresión:

$$P[\hat{p} = k] = P[X = nk].$$

Además la media y la varianza de \hat{p} son

$$\begin{aligned}\mu_{\hat{p}} &= E[\hat{p}] = E\left[\frac{X}{n}\right] = \frac{E[X]}{n} = \frac{np}{n} = p \\ \sigma_{\hat{p}}^2 &= Var[\hat{p}] = Var\left[\frac{X}{n}\right] = \frac{Var[X]}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n},\end{aligned}$$

respectivamente. Si n es grande, la distribución Binomial puede aproximarse por la distribución Normal, de modo que el estimador de p , $\hat{p} = \frac{X}{n}$, se distribuye según una Normal de parámetros p y $\sqrt{\frac{p(1-p)}{n}}$.

Este resultado es la base para desarrollar un gráfico de control de la fracción de disconformes.

Si se conoce el verdadero valor de la fracción disconforme p en el proceso de fabricación o se dispone de un valor estándar para ella, entonces, a partir del modelo general para el gráfico de control², la línea central y los límites de control del gráfico de la fracción disconforme serán:

$$\begin{aligned}LSC &= p + k\sqrt{\frac{p(1-p)}{n}} \\ LC &= p \\ LIC &= p - k\sqrt{\frac{p(1-p)}{n}}.\end{aligned}$$

²Se toma $w = p$ como estadístico para establecer los límites de control

Nota 30 Si los cálculos dan lugar a un LIC negativo, se tomará igual a 0, puesto que la fracción disconforme nunca puede tomar valores negativos.

Para utilizar el gráfico de control, se observan m muestras sucesivas de tamaño n del artículo inspeccionado, se calcula la fracción disconforme en cada muestra p_i , y se representa en el gráfico. Mientras que los p_i estén dentro de los límites de control y la sucesión de puntos no exhiba un patrón sistemático, se concluirá que el proceso está bajo control estadístico; mientras que si hay valores fuera de los límites de control o si se observa un patrón no aleatorio entre los puntos, habrá que concluir que la fracción de disconformes del proceso cambió hacia un nuevo nivel y el proceso está fuera de control.

También puede ocurrir que las m muestras consideradas tengan distintos tamaños n_i ($i = 1, \dots, m$). En este caso, cada muestra tiene sus propios límites de control relacionados con su tamaño, es decir:

$$\begin{aligned} LSC &= p + k\sqrt{\frac{p(1-p)}{n_i}} \\ LC &= p \\ LIC &= p - k\sqrt{\frac{p(1-p)}{n_i}} \end{aligned}$$

Para que todas las muestras tengan los mismos límites de control, se puede realizar el gráfico a partir de un tamaño muestral promedio, obteniendo así unos *límites de control constantes aproximados*. El tamaño muestral promedio es:

$$\bar{n} = \frac{1}{m} \sum_{i=1}^m n_i$$

y los límites de control aproximados son:

$$\begin{aligned} LSC &= p + 3\sqrt{\frac{p(1-p)}{\bar{n}}} \\ LC &= p \\ LIC &= p - 3\sqrt{\frac{p(1-p)}{\bar{n}}} \end{aligned}$$

Nota 31 *Si hay alguna muestra con valor muy próximo a los límites de control es conveniente calcular su límite exacto para ver si el proceso está bajo control.*

Nota 32 *La consideración de límites aproximados solo tiene sentido cuando los distintos tamaños muestrales están razonablemente próximos. Se considera que dichos tamaños muestrales están próximos si todos ellos se encuentran en torno al 25% a cada lado del tamaño muestral promedio, esto es, si $n_i \in (\frac{3}{4}\bar{n}, \frac{5}{4}\bar{n})$ para todo $i = 1, \dots, m$.*

Cuando se desconoce la fracción disconforme p de artículos del proceso, hay que estimarla a partir de los datos observados³. En este caso, los límites de control obtenidos se pueden considerar límites de control de prueba (en el periodo base) y estos límites permiten comprobar si el proceso estaba bajo control cuando se obtuvieron las m muestras iniciales. Para ello, se representa en el gráfico la fracción muestral disconforme para cada muestra p_i y se compara con los límites de control obtenidos. Si todos los puntos están dentro de los límites de control y no se manifiesta un comportamiento sistemático, entonces se concluirá que el proceso estaba bajo control en el pasado, y que los límites de control de prueba son adecuados para controlar la producción actual y futura.

Si se quiere controlar la producción actual o futura, hay que basarse en datos de un proceso que está bajo control. Por el contrario, si uno o más de los puntos p_i se encuentran fuera de los límites de control de prueba, se rechaza la hipótesis de un estado de control anterior, por lo que es necesario revisar los límites de control de prueba. Para ello, se examina cada uno de los puntos obtenidos fuera de control y se busca una causa atribuible a cada uno de ellos. Si se encuentra la causa por la que la muestra estaba fuera de los límites de control, se descarta la muestra del estudio y se vuelven a calcular los límites de control de prueba, utilizando únicamente las muestras restantes. Entonces, se representa el nuevo gráfico de control y se examina si las muestras restantes están bajo control con estos nuevos límites (los puntos que inicialmente estaban bajo control podrían salir fuera de estos nuevos límites, ya que éstos, a veces, son más estrechos que los anteriores). Este proceso continúa hasta que todas las muestras estén bajo control y, a partir de este momento, se aceptan los límites de control de prueba como límites de control para vigilar el proceso (para el periodo de vigilancia).

³Posteriormente se verá cómo.

Si el gráfico de control se basa en un valor conocido o estándar de la fracción disconforme p , entonces el cálculo de límites de control de prueba suele ser innecesario. Sin embargo, se debe tener cuidado al trabajar con un valor estándar de p , puesto que en la práctica raramente se conoce con exactitud el verdadero valor de p , y normalmente se proporciona un valor estándar que representa un valor deseado u objetivo para la fracción disconforme del proceso. Si está fuera de control, se tendría que determinar si el proceso está fuera de control para ese p , pero dentro de control para otro valor de p . Por ejemplo, si se especifica un valor objetivo de $p = 0.18$, pero el proceso realmente está bajo control para un valor de la fracción disconforme $p = 0.15$. Utilizando el diagrama de control con base $p = 0.18$, muchos puntos caerán por debajo del límite inferior de control, indicando una situación de fuera de control. Sin embargo, el proceso sólo está fuera de control respecto al objetivo $p = 0.18$. Algunas veces es posible mejorar el nivel de calidad usando valores objetivos, o poniendo bajo control un proceso a un nivel particular de calidad. Los valores objetivo para p pueden ser útiles en procesos en los cuales es posible controlar la fracción disconforme mediante ajustes sencillos.

Nota 33 *Estas observaciones son válidas también para el resto de gráficos de control de atributos.*

Como ya se ha indicado, cuando se desconoce el valor de p es necesario estimarlo a partir de las m muestras extraídas durante el periodo base. Sin embargo, el número de artículos contenidos en cada muestra puede ser variable, por lo que distinguiremos los casos de tamaño muestral constante y variable.

3.2.1 Tamaño Muestral Constante

Para estimar el valor del parámetro desconocido en el periodo base, se toman m muestras preliminares todas ellas de tamaño n . Como regla general, m tendría que ser al menos 20. Entonces, si hay x_i artículos disconformes en la muestra i , se calcula la fracción disconforme de la i -ésima muestra, \hat{p}_i , como:

$$\hat{p}_i = \frac{x_i}{n} \quad i = 1, 2, \dots, m$$

y la media de estas fracciones disconformes muestrales individuales es:

$$\bar{p} = \frac{1}{m} \sum_{i=1}^m \hat{p}_i = \frac{1}{m \cdot n} \sum_{i=1}^m x_i.$$

Este estadístico \bar{p} estima la fracción disconforme p desconocida, por lo que los límites de control del gráfico de control de la fracción disconforme son los siguientes:

$$\begin{aligned} LSC &= \bar{p} + k\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \\ LC &= \bar{p} \\ LIC &= \bar{p} - k\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}. \end{aligned}$$

Ejemplo 34 *Los resultados de muestras obtenidas para la inspección de un cierto producto durante los primeros 25 días del mes de febrero aparecen en la Tabla 3.1. En ella se ha calculado la proporción de artículos disconformes para cada día.*

Si se consideran estos 25 días de febrero como período base, la estimación de p es

$$\bar{p} = \frac{\sum_{i=1}^m x_i}{n \cdot m} = \frac{47}{5000} = \frac{\sum_{i=1}^m \hat{p}_i}{m} = \frac{0.235}{25} = 0.0094.$$

Los límites de control están representados en la Figura 3.1 y son, por tanto:

$$\begin{aligned} LSC &= 0.0094 + 3\sqrt{\frac{0.0094(1-0.0094)}{200}} = 0.0299 \\ LC &= 0.0094 \\ LIC &= 0.0094 - 3\sqrt{\frac{0.0094(1-0.0094)}{200}} = 0^4. \end{aligned}$$

Si se revisa la columna de la fracción defectuosa, se ve que las muestras correspondientes a los días 4 y 9 de febrero están fuera de los límites de control. De estos dos días, se puede llegar a establecer la causa del mayor número de unidades defectuosas producidas el 9 de febrero. Si no parece

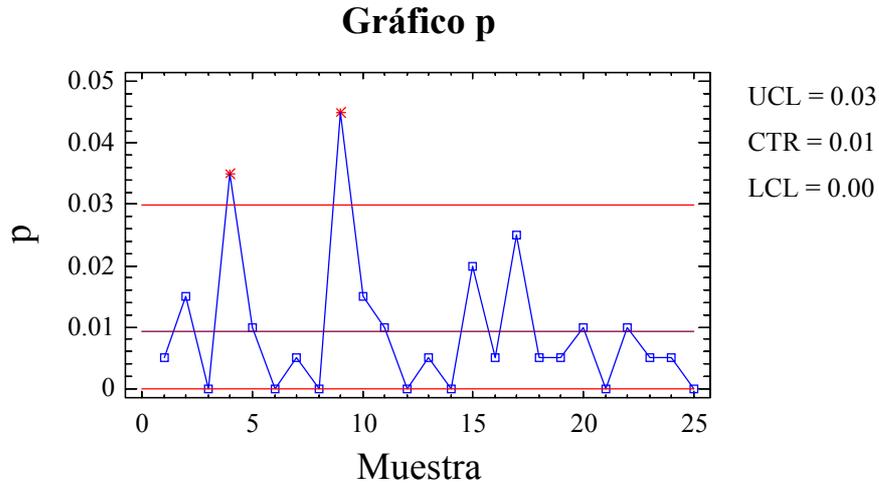


Figura 3.1: Gráfico p para los 25 primeros días de febrero

probable que se repita, se elimina el punto del análisis del período base. El objetivo del análisis del período base es establecer qué estándar se puede lograr. La causa del punto fuera de control que se produjo el 4 de febrero no se puede establecer, de modo que se conserva en el análisis como parte de la variabilidad natural del proceso. Una vez excluido el lote del 9 de febrero, la estimación de p y los nuevos límites de control están representados en la Figura 3.2 y son:

$$\begin{aligned}\bar{p} &= \frac{38}{4800} = 0.0079 \\ LSC &= 0.0267 \\ LC &= 0.0079 \\ LIC &= 0.\end{aligned}$$

Ahora el punto correspondiente al día 4 de febrero está fuera de los límites de control. Como es la única muestra que está fuera de control, se ve lo que sucede si se excluye del análisis, de manera que los nuevos límites

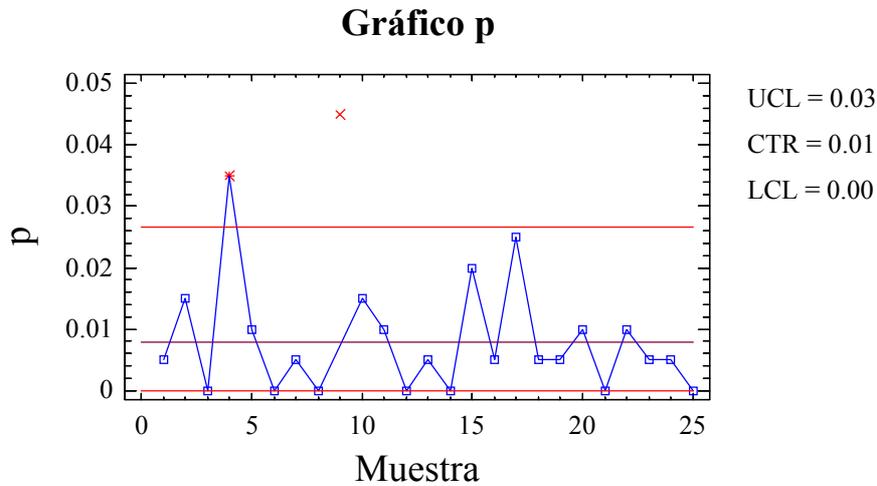


Figura 3.2: Gráfico p tras excluir el 9 de febrero

de control están representados en la Figura 3.3 y son:

$$\begin{aligned}\bar{p} &= \frac{31}{4600} = 0.0067 \\ LSC &= 0.0241 \\ LC &= 0.0067 \\ LIC &= 0\end{aligned}$$

Entonces, el punto asociado al día 17 de febrero se sale de los límites de control, por lo que sería conveniente introducir de nuevo el correspondiente al día 4 de febrero en el análisis y seguir considerando su anomalía como una situación normal de causa aleatoria.

3.2.2 Tamaño Muestral Variable

En algunos casos, se pueden producir diferentes cantidades de artículos en cada período de inspección, con lo que el gráfico de control tendría entonces un tamaño muestral variable.

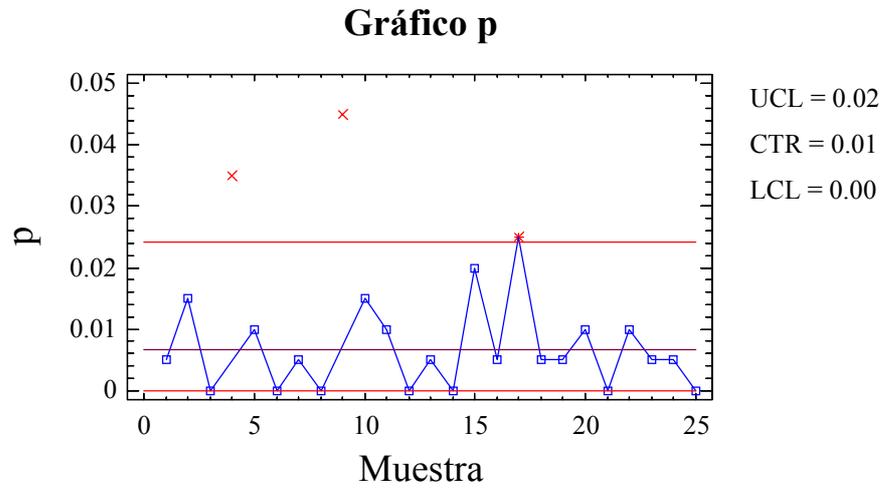


Figura 3.3: Gráfico p tras excluir 4 y 9 de febrero

Si se tienen m muestras, cada una de ellas de tamaño n_i , $i = 1, \dots, m$, donde \hat{p}_i es la fracción de disconformes en la muestra i -ésima, la estimación de p en el periodo base es el promedio ponderado:

$$\bar{p} = \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m n_i} = \frac{\sum_{i=1}^m n_i \hat{p}_i}{\sum_{i=1}^m n_i}$$

que es distinto del valor $\frac{1}{m} \sum_{i=1}^m \hat{p}_i$.

Por tanto, los límites de control para el período i –ésimo son:

$$\begin{aligned} LSC &= \bar{p} + k\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}} \\ LC &= \bar{p} \\ LIC &= \bar{p} - k\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}} \end{aligned}$$

es decir, para cada muestra individual se tienen los límites de control basados en su tamaño muestral, también denominados *límites específicos*.

Ejemplo 35 *Los resultados de muestras obtenidas para la inspección un cierto producto durante los primeros 25 días del mes de febrero aparecen en la Tabla 3.2. En ella, se han añadido los límites de control propios de cada muestra.*

El estimador de la fracción de disconformes, \bar{p} para las 25 muestras es:

$$\bar{p} = \frac{\sum_{i=1}^{25} x_i}{\sum_{i=1}^{25} n_i} = \frac{25}{249} = 0.1004$$

Entonces, los límites de control para cada muestra se obtienen como:

$$\begin{aligned} LSC &= 0.1004 + 3\sqrt{\frac{0.1004(1-0.1004)}{n_i}} \\ LC &= 0.1004 \\ LIC &= 0.1004 - 3\sqrt{\frac{0.1004(1-0.1004)}{n_i}} \end{aligned}$$

y se recogen en la Tabla 3.2, junto con la fracción disconforme de cada muestra. Además, están representados en la Figura 3.4 y se puede observar que el proceso está bajo control estadístico en el periodo base.

Un método alternativo a éste, para obtener unos límites de control iguales para todas las muestras, consiste en realizar el gráfico de control

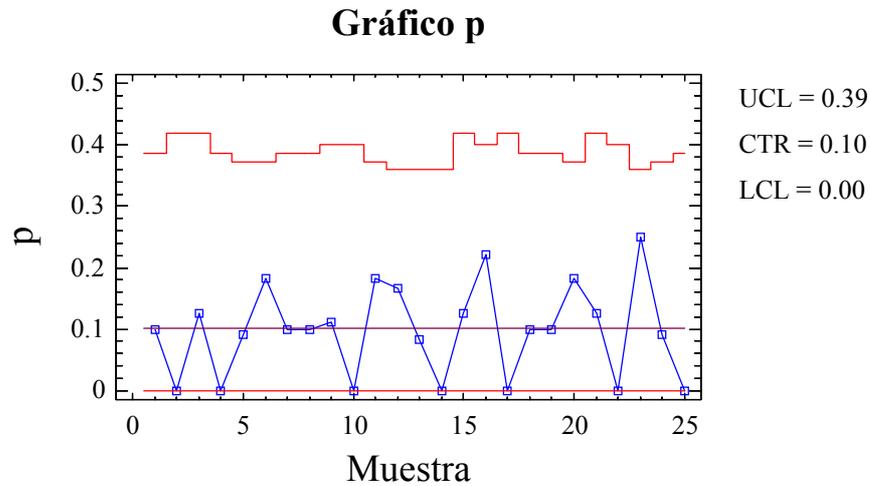


Figura 3.4: Gráfico p de tamaño muestral variable para los 25 días de febrero

a partir del tamaño muestral promedio, obteniendo así unos *límites de control aproximados*, que se calculan como⁵:

$$\begin{aligned} LSC &= \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}} \\ LC &= \bar{p} \\ LIC &= \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}. \end{aligned}$$

Ejemplo 36 En el Ejemplo 2, el tamaño muestral promedio es

$$\bar{n} = \frac{249}{25} = 9.96$$

y el intervalo de tamaños muestrales que da su 25% es (7.47, 12.45), con lo que todos los tamaños muestrales están dentro del intervalo y se puede utilizar el tamaño muestral promedio.

⁵En estos límites de control hay que considerar las Notas 3 y 4

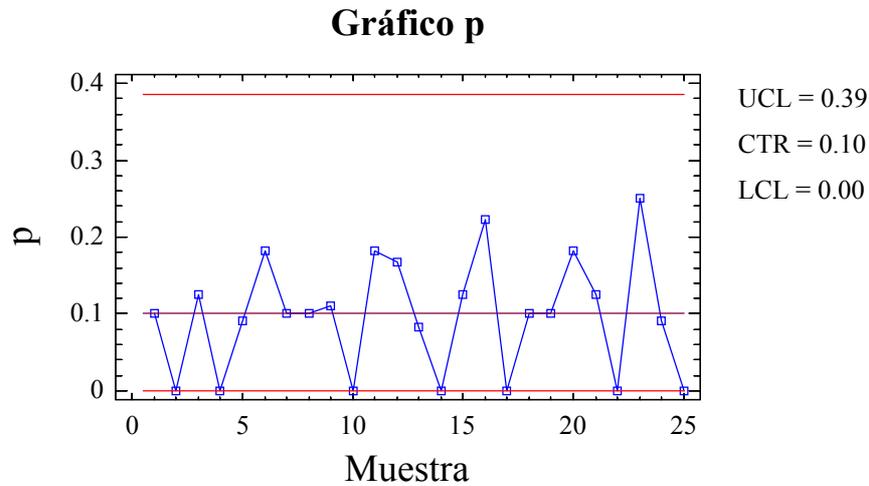


Figura 3.5: Gráfico p de tamaño muestral promedio para los 25 días de febrero

Entonces, los límites de control aproximados están representados en la Figura 3.5 y son:

$$LSC = 0.1004 + 3\sqrt{\frac{0.1004(1 - 0.1004)}{9.96}} = 0.3861$$

$$LC = 0.1004$$

$$LIC = 0.1004 - 3\sqrt{\frac{0.1004(1 - 0.1004)}{9.96}} = 0.$$

3.2.3 Gráfico p Normalizado

Este gráfico consiste en la normalización del gráfico p , de manera que los límites de control estén en los valores $\pm k$ y la línea central en cero. Para ello, se representan en el gráfico los puntos:

$$z_i = \frac{p_i - \hat{p}}{\sqrt{\frac{\hat{p}(1 - \hat{p})}{n_i}}}$$

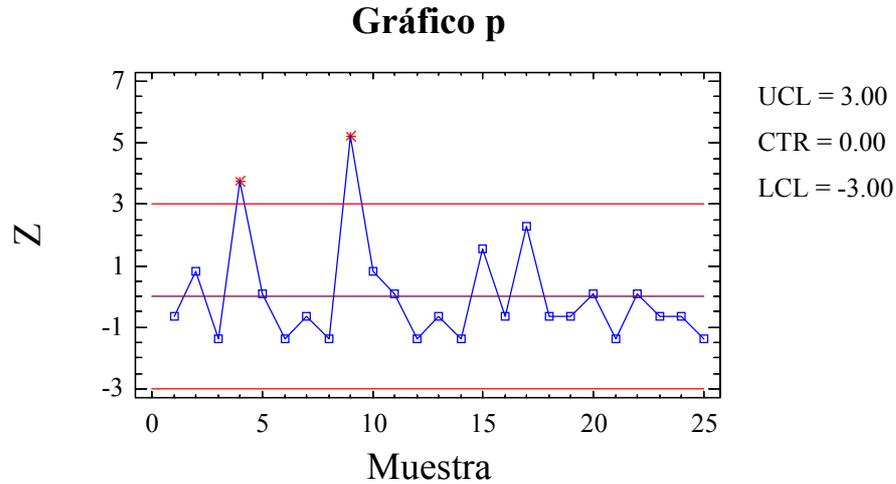


Figura 3.6: Gráfico p normalizado del Ejemplo 35

Ejemplo 37 Los valores de p normalizados y los gráficos p normalizados de los Ejemplos 35 y 37 se encuentran en las Tablas 3.1 y 3.2 y en las Figuras 3.6 y 3.7, respectivamente.

3.3 Gráfico np

El gráfico np , que estudia el número de artículos disconformes, se reduce directamente al gráfico p con todos los tamaños muestrales iguales que se ha estudiado, sin más que multiplicar por n los límites de control obtenidos.

Entonces, si se conoce el verdadero valor del número de artículos disconformes np (o bien de la fracción disconforme p) en el proceso de fabricación, o se dispone de un valor estándar para él, los límites de control para el gráfico np son:

$$\begin{aligned} LSC &= np + k\sqrt{np(1-p)} \\ LC &= np \\ LIC &= np - k\sqrt{np(1-p)}. \end{aligned}$$

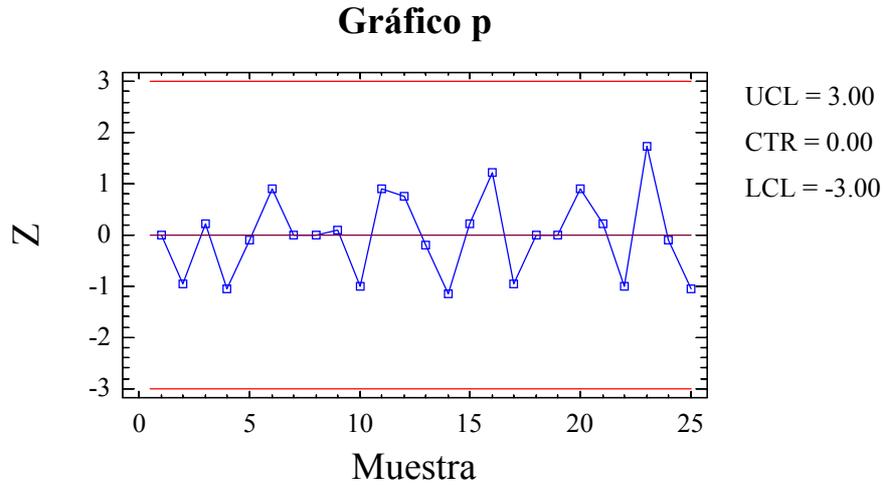


Figura 3.7: Gráfico p normalizado del Ejemplo 37

En el caso en el que se desconoce el verdadero valor del número de artículos disconformes np (o la fracción disconforme p), hay que estimarla a partir de los datos observados. Para ello, se toman m muestras preliminares, todas ellas de tamaño n , en las que hay x_i artículos no conformes en la muestra i . Entonces, la media de artículos disconformes es:

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i = n\bar{p}.$$

Este estadístico \bar{x} estima el número de artículos disconformes y de él se puede obtener el valor de $\bar{p} = \frac{\bar{x}}{n}$ que se va a utilizar en los límites de control. Entonces, los límites de control del gráfico de control de artículos disconformes son los siguientes:

$$\begin{aligned} LSC &= n\bar{p} + k\sqrt{n\bar{p}(1-\bar{p})} \\ LC &= n\bar{p} \\ LIC &= n\bar{p} - k\sqrt{n\bar{p}(1-\bar{p})}. \end{aligned}$$

Ejemplo 38 Se retoman los datos del Ejemplo 35 (véase Tabla 3.1) y se construye un gráfico np .

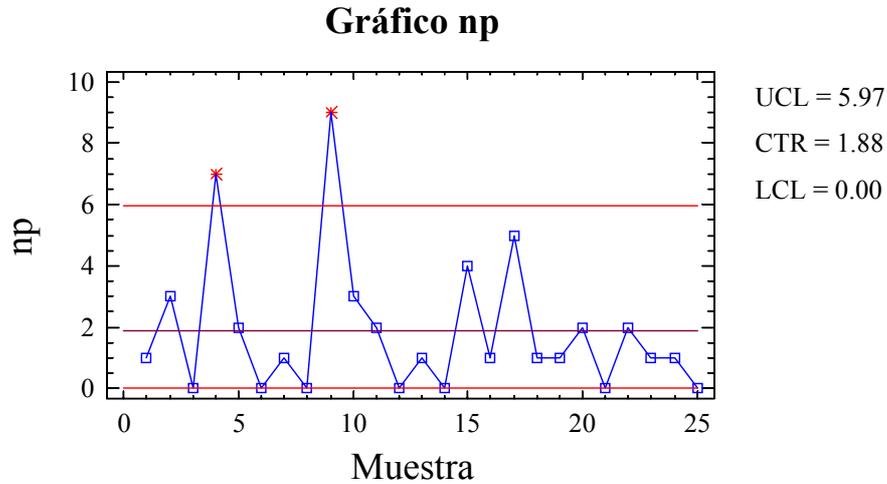


Figura 3.8: Gráfico np para los primeros 25 días de febrero

Si se consideran los artículos inspeccionados durante los primeros 25 días de febrero como período base, la estimación de los parámetros necesarios es:

$$\bar{x} = \frac{\sum_{i=1}^m x_i}{m} = \frac{47}{25} = 1.88$$

$$\bar{p} = \frac{\bar{x}}{n} = 0.0094.$$

Los límites de control están representados en la Figura 3.8 y son, por tanto:

$$LSC = 1.88 + 3\sqrt{1.88(1 - 0.0094)} = 5.9740$$

$$LC = 1.88$$

$$LIC = 1.88 - 3\sqrt{1.88(1 - 0.0094)} = 0.$$

El análisis es totalmente análogo al caso del gráfico p , así si se revisa el número de artículos disconformes en cada día, se puede ver que los de los días 4 y 9 de febrero están fuera de los límites de control. De estos dos días, se puede llegar a establecer la causa del mayor número de unidades

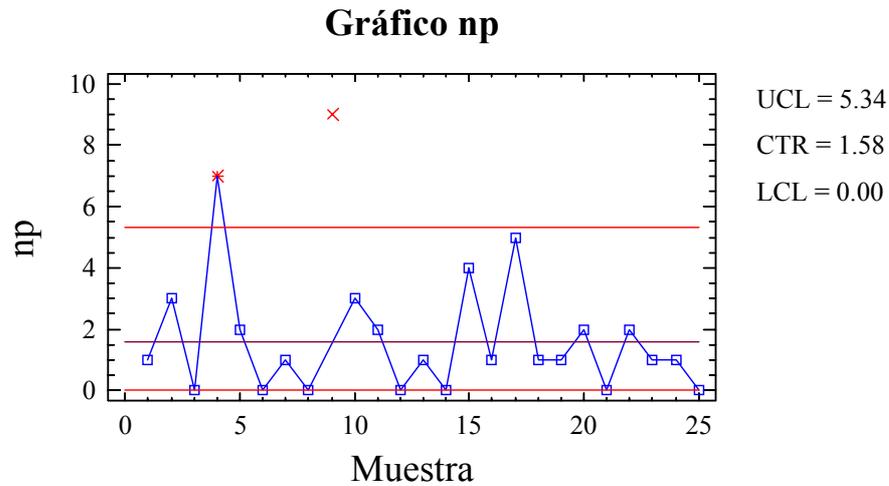


Figura 3.9: Gráfico np tras excluir el 4 de febrero

defectuosas producidas el 9 de febrero. Si no parece probable que se repita, se elimina el punto del análisis del período base. El objetivo del análisis del período base es establecer qué estándar se puede lograr. La causa del punto fuera de control que se produjo el 4 de febrero no se puede establecer, de modo que se conserva en el análisis como parte de la variabilidad natural del proceso. Una vez excluida la producción inspeccionada el 9 de febrero, la estimación de p y los nuevos límites de control, representados en la Figura 3.9, son:

$$\begin{aligned}\bar{x} &= 1.5833 \\ \bar{p} &= 0.0079 \\ LSC &= 5.3433 \\ LC &= 1.5833 \\ LIC &= 0.\end{aligned}$$

Ahora el punto correspondiente al día 4 de febrero está fuera de los límites de control. Si se excluye del análisis, los límites de control, representados

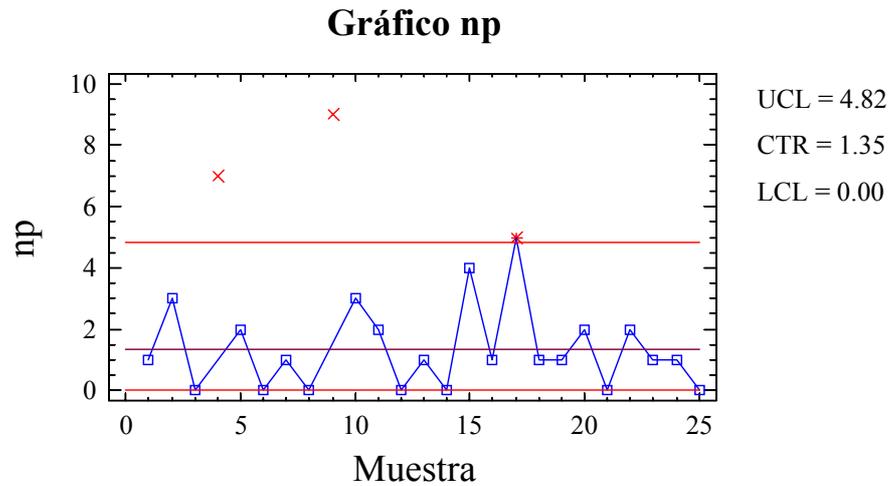


Figura 3.10: Gráfico np tras excluir los días 4 y 9 de febrero

en la Figura 3.10, son:

$$\begin{aligned}\bar{x} &= 1.3478 \\ \bar{p} &= 0.0067 \\ LSC &= 4.8189 \\ LC &= 1.3478 \\ LIC &= 0\end{aligned}$$

Entonces, la muestra del día 17 de febrero se sale de los límites de control, por lo que sería conveniente introducir de nuevo la del día 4 en el análisis y seguir considerando su anomalía como una situación normal de causa aleatoria.

Nota 39 Se puede observar que los gráficos p y np son idénticos salvo la escala del eje vertical.

3.4 Gráficos de Control de Disconformidades

Cada punto concreto en el que no se satisface una especificación es un defecto o disconformidad; por tanto, un artículo disconforme tendrá, por lo menos, una disconformidad, aunque dependiendo de su naturaleza y su gravedad, es posible que posea varias disconformidades y no sea clasificado como tal, siempre que su funcionamiento no se vea afectado seriamente. Hay muchos casos en los que es preferible trabajar directamente con el conjunto de disconformidades en lugar de la fracción de disconformes, por ejemplo si se estudia el número de parches en 50 metros de tubería. La condición para el uso de este nuevo gráfico es que la unidad de inspección sea homogénea en todas las muestras. Las unidades de inspección no tienen que tener el mismo número de artículos, ya que, por ejemplo, se inspeccionan cajas de alfileres con un contenido similar, porque revisar los alfileres uno a uno es una labor tediosa y costosa. Así se pueden crear dos tipos de gráficos de control:

1. *Gráfico del número de disconformidades en una unidad de inspección o Gráfico c* , cuando la unidad de inspección permanece razonablemente constante de un período a otro.
2. *Gráfico del número de disconformidades por unidad o Gráfico u* , cuando se quiere expresar el número total de disconformidades respecto a un número de unidades de inspección n .

3.4.1 Gráfico de Control del Número de Disconformidades en una Unidad de Inspección: Gráfico c

Se estudia el número de disconformidades en una unidad de inspección con tamaño muestral constante del artículo observado, siendo una unidad de inspección una entidad apropiada para registrar las disconformidades (por ejemplo, un número concreto de artículos). Si se define la variable aleatoria C como el *número de disconformidades por unidad de inspección*, ésta se distribuye según una distribución de Poisson de parámetro c ($c > 0$), esto es:

$$P[C = x] = e^{-c} \frac{c^x}{x!}, \quad x = 0, 1, 2, \dots$$

La media y varianza de la distribución de Poisson son iguales y coinciden con el valor del parámetro, c . Por tanto, los límites de control para el gráfico c con los límites de $k\sigma$ son:

$$\begin{aligned}LSC &= c + k\sqrt{c} \\LC &= c \\LIC &= c - k\sqrt{c}\end{aligned}$$

supuesto que se tiene un valor estándar para el parámetro desconocido c .

Nota 40 Si los cálculos dan lugar a un LIC negativo, se tomará igual a 0.

Si no se tiene un valor estándar para c , se estimará mediante el número medio de disconformidades presentes en las m muestras tomadas en el período base:

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

Los límites de control serán, entonces:

$$\begin{aligned}LSC &= \bar{c} + k\sqrt{\bar{c}} \\LC &= \bar{c} \\LIC &= \bar{c} - k\sqrt{\bar{c}}\end{aligned}$$

Ejemplo 41 Los resultados obtenidos al observar el número de disconformidades en muestras de 200 cristales para ventanas se encuentran en la Tabla 3.3.

La unidad de inspección está formada por 200 artículos. El número de disconformidades es de 410 y hay 22 muestras, así que la estimación de c es $\bar{c} = \frac{410}{22} = 18.6364$. Los límites de control 3σ para el período base se representan en la Figura 3.11 y vienen dados por:

$$\begin{aligned}LSC &= \bar{c} + 3\sqrt{\bar{c}} = 18.6364 + 3\sqrt{18.6364} = 31.5873 \\LC &= \bar{c} = 18.6364 \\LIC &= \bar{c} - 3\sqrt{\bar{c}} = 18.6364 - 3\sqrt{18.6364} = 5.6854.\end{aligned}$$

Ningún punto sale fuera de los límites de control de modo que éstos serán los valores estándares de los límites de control para el período de vigilancia.

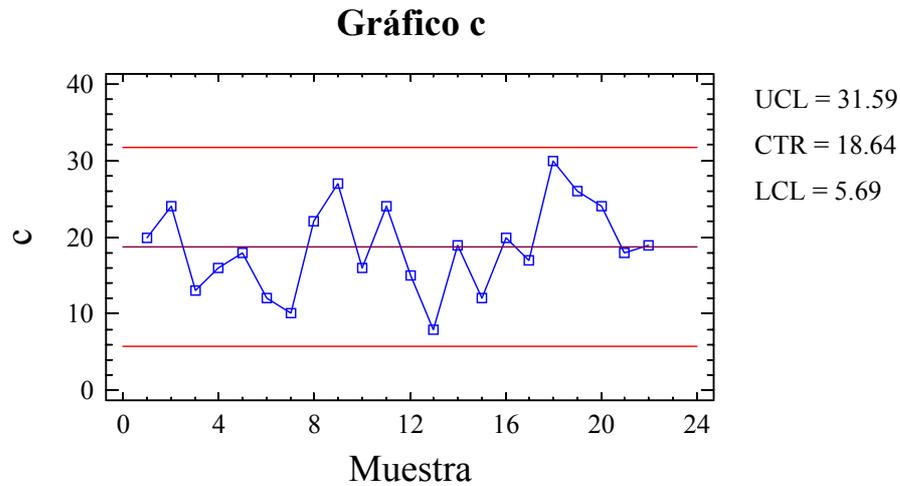


Figura 3.11: Gráfico c en 22 muestras de 200 cristales

3.4.2 Gráfico c Normalizado

Este gráfico consiste en la normalización del gráfico c , de manera que los límites de control estén situados en los valores $\pm k$ y la línea central en cero. Para ello, se representan en el gráfico los puntos:

$$z_i = \frac{c_i - \hat{c}}{\sqrt{\hat{c}}}$$

Ejemplo 42 En la Tabla 3.3 se encuentran los valores normalizados y en la Figura 3.12 se puede observar el gráfico c normalizado del Ejemplo 41. Obsévese que los gráficos c y c normalizado solo difieren en la escala del eje Y : su trazo es idéntico.

3.4.3 Gráfico de Control del Número de Disconformidades por Unidad: Gráfico u

Para llevar a cabo el estudio de este gráfico se distinguen dos casos, atendiendo al tamaño muestral, que puede ser constante o variable.

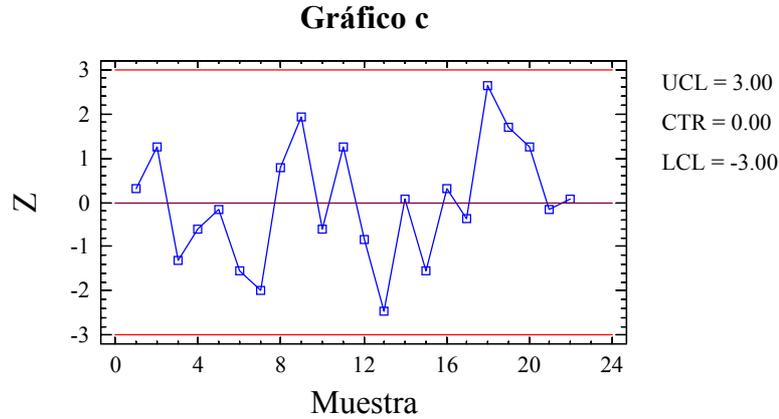


Figura 3.12: Gráfico c normalizado del Ejemplo 41

Tamaño Muestral Constante

Si hay X disconformidades en una muestra de n unidades, donde la variable aleatoria C sigue una distribución de Poisson de parámetro c , entonces el número de disconformidades por unidad será:

$$U = \frac{C}{n}.$$

Se sabe que U es una variable aleatoria de Poisson, puesto que es una combinación lineal de variables aleatorias independientes de Poisson, y su distribución de probabilidad está dada por:

$$P[U \leq a] = P[C \leq na] = \sum_{x=0}^{[na]} \frac{e^{-c} (c)^x}{x!}$$

donde $[na]$ es el mayor entero menor o igual que na . La media y la varianza de U son:

$$\begin{aligned} \mu_U &= E[U] = E\left[\frac{C}{n}\right] = \frac{c}{n} = u \\ \sigma_U^2 &= Var[U] = Var\left[\frac{C}{n}\right] = \frac{c}{n^2} = \frac{u}{n}. \end{aligned}$$

Por tanto, los límites de control para el gráfico u con los límites de $k\sigma$ se calculan como:

$$\begin{aligned}LSC &= u + k\sqrt{\frac{u}{n}} \\LC &= u \\LIC &= u - k\sqrt{\frac{u}{n}}\end{aligned}$$

si se tiene un valor estándar para el parámetro desconocido u .

Nota 43 Si los cálculos dan lugar a un LIC negativo, se tomará igual a 0.

Si no se dispone de un valor estándar para u , se estimará mediante el número de disconformidades por unidad presentes en las m muestras tomadas en el período base:

$$\bar{u} = \frac{1}{n \cdot m} \sum_{i=1}^m c_i = \frac{1}{m} \sum_{i=1}^m u_i$$

Por tanto, los límites para el gráfico de control son:

$$\begin{aligned}LSC &= \bar{u} + k\sqrt{\frac{\bar{u}}{n}} \\LC &= \bar{u} \\LIC &= \bar{u} - k\sqrt{\frac{\bar{u}}{n}}.\end{aligned}$$

Ejemplo 44 Un fabricante de lavadoras desea establecer un gráfico de control de disconformidades por unidad en la línea de montaje final. Los resultados de 20 muestras de cinco lavadoras cada una aparecen en la Tabla 3.4.

$$\text{El valor } \bar{u} = \frac{211}{100} = \frac{42.2}{20} = 2.11$$

Los límites de control 3σ quedan:

$$\begin{aligned}LSC &= \bar{u} + 3\sqrt{\frac{\bar{u}}{n}} = 2.11 + 3\sqrt{\frac{2.11}{5}} = 4.058 \\LC &= \bar{u} = 2.11 \\LIC &= \bar{u} - 3\sqrt{\frac{\bar{u}}{n}} = 2.11 - 3\sqrt{\frac{2.11}{5}} = 0\end{aligned}$$

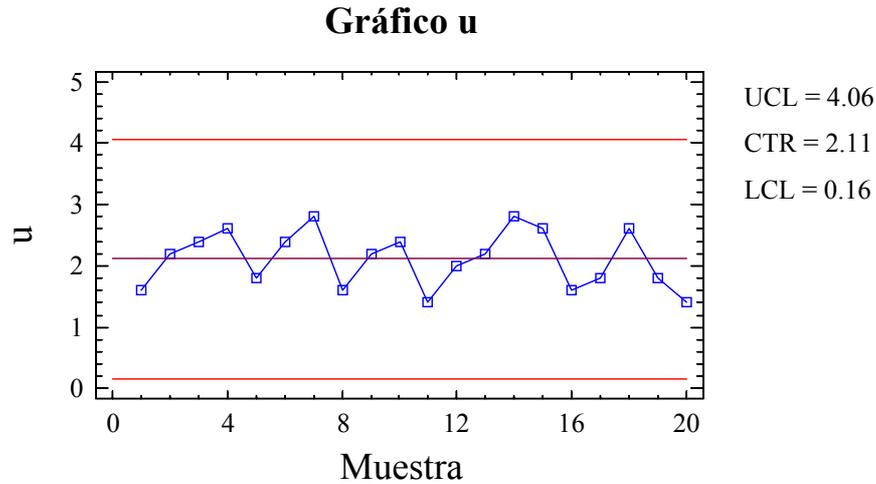


Figura 3.13: Gráfico u para 20 muestras de 5 lavadoras

y se representan en la Figura 3.13. Estos datos no presentan una falta de control estadístico; por lo tanto, se adoptarían estos límites para el período de vigilancia. Así, un estándar realista para el número medio de disconformidades en la línea de montaje final por lavadora será $\bar{u} = 2.11$.

Tamaño Muestral Variable

En muchos casos el número de unidades en una muestra no es constante, por ejemplo, en la inspección del número de erratas en documentos. El tamaño muestral varía, pues no todos los documentos tienen la misma extensión. Entonces, se debe utilizar un gráfico de control del número de disconformidades por unidad (gráfico u) que tendrá una línea central constante, pero los límites de control variarán en forma inversamente proporcional al tamaño de la muestra.

Por tanto, los límites de control para el gráfico u en la muestra i –sima

con los límites de $k\sigma$ son:

$$\begin{aligned}LSC &= u + k\sqrt{\frac{u}{n_i}} \\LC &= u \\LIC &= u - k\sqrt{\frac{u}{n_i}}\end{aligned}$$

dado un valor estándar para el parámetro desconocido u .

Si no se tiene un valor estándar para u , se estimará mediante el número de disconformidades por unidad presentes en las m muestras tomadas en el período base, o lo que es lo mismo, el valor ponderado de cada u_i , denotado como \bar{u} y que se calcula como:

$$= \bar{u} = \frac{\sum_{i=1}^m c_i}{\sum_{i=1}^m n_i} = \frac{\sum_{i=1}^m n_i u_i}{\sum_{i=1}^m n_i}.$$

Por tanto, los límites para el gráfico de control en la muestra i – *sima* vienen dados por:

$$\begin{aligned}LIC &= \bar{u} - k\sqrt{\frac{\bar{u}}{n_i}} \\LC &= \bar{u} \\LSC &= \bar{u} + k\sqrt{\frac{\bar{u}}{n_i}}.\end{aligned}$$

Nota 45 *En este caso, también se puede trabajar con el tamaño muestral promedio, en las mismas condiciones que se comentaron para el gráfico p.*

Ejemplo 46 *En una empresa se inspeccionan libros para descubrir el número de erratas, considerando cada capítulo como una unidad de inspección. El valor estándar del número de erratas por capítulo es de 1.5 y se quiere saber si este valor estándar se mantiene en vigor. Los datos para diez libros aparecen en la Tabla 3.5.*

Como se conoce el valor estándar de $u = 1.5$, se trabaja en el periodo de vigilancia. Por consiguiente, los límites de control 3σ para el periodo

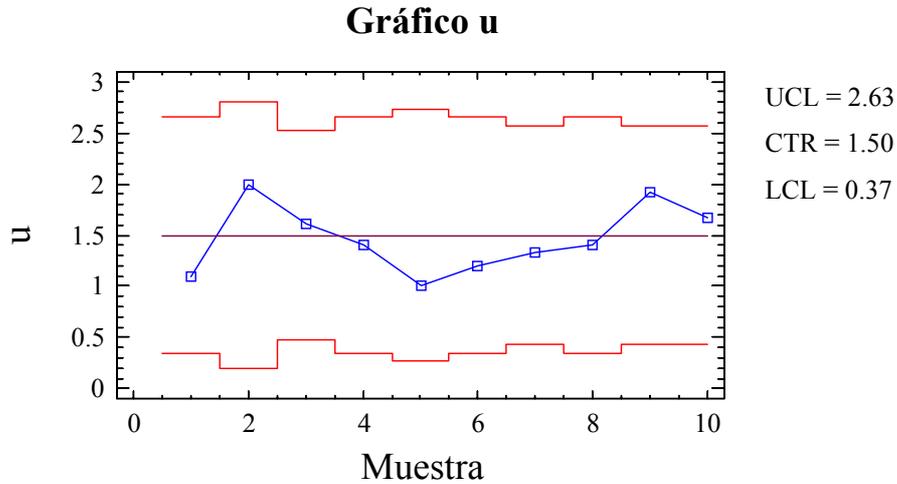


Figura 3.14: Gráfico u del número de erratas por capítulo

de vigilancia son:

$$LSC = u + 3\sqrt{\frac{u}{n_i}} = 1.5 + 3\sqrt{\frac{1.5}{n_i}}$$

$$LC = u = 1.5$$

$$LIC = u - 3\sqrt{\frac{u}{n_i}} = 1.5 - 3\sqrt{\frac{1.5}{n_i}}.$$

Así, los límites de control para cada tamaño muestral se encuentran en la Tabla 3.6.

Se puede observar el gráfico u en la Figura 3.14, donde se ve que el proceso está bajo control estadístico en el periodo de vigilancia, pues todas las muestras se encuentran dentro de los límites de control.

Gráfico u Normalizado

Este gráfico consiste en la normalización del gráfico u , de manera que los límites de control estén en los valores $\pm k$ y la línea central en cero. Para

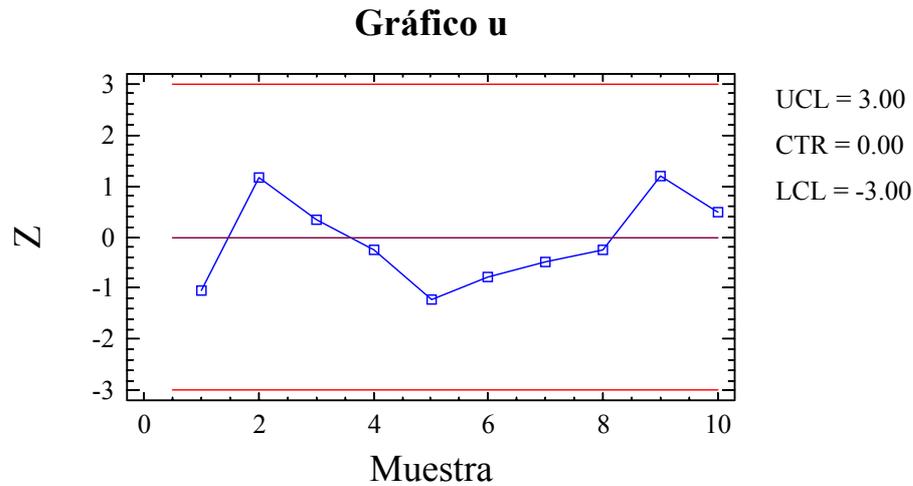


Figura 3.15: Gráfico u normalizado del Ejemplo 46

ello, se representan en el gráfico los puntos:

$$z_i = \frac{u_i - \bar{u}}{\sqrt{\frac{\bar{u}}{n_i}}}$$

Ejemplo 47 En la Tabla 3.6 se encuentran los valores normalizados y en la Figura 3.15 se puede ver el gráfico u normalizado del Ejemplo 46.

3.5 Sistema de Calificación de Deméritos

Si se estudian productos complejos como electrodomésticos, vehículos a motor o aparatos mayores, se puede comprobar que pueden tener muchos tipos de disconformidades o defectos y no todos ellos de la misma importancia. Probablemente un artículo con un defecto muy grave se clasificaría como disconforme respecto a los requisitos, pero uno con varios defectos menores no necesariamente tendría que ser disconforme. En estos casos es necesario un método que permita clasificar las disconformidades o defectos

de acuerdo con su gravedad y ponderar los diversos tipos de defectos de un modo razonable. Una clasificación de las disconformidades es:

- **Clase A (defectos muy graves):** El artículo es completamente inadecuado para su utilización o fallará de manera que no se pueda reparar con facilidad en el lugar de trabajo, o bien ocasionará lesiones personales o daños materiales.
- **Clase B (defectos graves):** El artículo quizás sufrirá un fallo del tipo A, generará seguramente problemas operacionales menos graves o reducirá su duración o incrementará los costos de mantenimiento.
- **Clase C (defectos moderadamente importantes):** El artículo probablemente fallará durante su utilización, causará problemas menos graves que los defectos anteriores, tendrá tal vez duración reducida, producirá un aumento en los costos de mantenimiento, o bien tendrá un defecto importante en el acabado, la presentación o la calidad del trabajo.
- **Clase D (defectos poco importantes):** El artículo no fallará durante su utilización, pero presenta defectos menores en el acabado, la presentación o la calidad del trabajo.

En el *plan Dodge-Torrey* los defectos se clasifican dentro de los grupos A, B, C y D, con un peso ponderado de 100, 50, 10 y 1, respectivamente, aunque se pueden considerar otros pesos más adecuados dependiendo del trabajo que se esté realizando.

Sean C_A , C_B , C_C y C_D variables aleatorias que miden el número de defectos de cada tipo en una unidad de inspección. Así:

$$C_K \rightsquigarrow P(c_K), \quad K = A, B, C, D$$

con c_K el número medio de defectos de tipo K .

Sea D la variable aleatoria que mide el número de deméritos en una unidad de inspección, definida como:

$$D = 100C_A + 50C_B + 10C_C + C_D.$$

Entonces, si la muestra contiene n unidades, la variable aleatoria $U = \frac{D}{n}$ mide el número de deméritos por unidad y sus parámetros son:

$$\begin{aligned}\mu_U &= E\left[\frac{D}{n}\right] = \frac{E[D]}{n} = \frac{100c_A + 50c_B + 10c_C + c_D}{n} = 100u_A + 50u_B + 10u_C + u_D \\ \sigma_U^2 &= Var\left[\frac{D}{n}\right] = \frac{Var[D]}{n^2} = \frac{100^2c_A + 50^2c_B + 10^2c_C + c_D}{n^2} = \\ &= \frac{100^2u_A + 50^2u_B + 10^2u_C + u_D}{n}\end{aligned}$$

donde $u_K = \frac{c_K}{n}$ es el número de defectos por unidad de tipo $K = A, B, C, D$.

En consecuencia, el gráfico de control para el número de deméritos por unidad cuando se conocen los valores estándares de u_A, u_B, u_C y u_D tiene los siguientes límites de control:

$$\begin{aligned}LSC &= \mu_U + k\sigma_U \\ LC &= \mu_U \\ LIC &= \mu_U - k\sigma_U.\end{aligned}$$

En caso contrario, se estiman dichos parámetros mediante el número medio de defectos de cada tipo presentes en las m muestras tomadas en el periodo base. Así, si se extraen m muestras de igual tamaño n , el número de deméritos por unidad en la muestra i -ésima viene dado por:

$$u_i = 100u_i^A + 50u_i^B + 10u_i^C + u_i^D$$

donde u_i^K es el número de defectos de tipo K por unidad en la muestra i , es decir, $\frac{c_i^K}{n}$ para $K = A, B, C, D$.

Por tanto, una estimación de la media y de la varianza del número de deméritos por unidad es:

$$\hat{\mu}_U = \bar{u} = \frac{1}{m} \sum_{i=1}^m u_i = 100\bar{u}_A + 50\bar{u}_B + 10\bar{u}_C + \bar{u}_D$$

$$\hat{\sigma}_U^2 = \frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{n}$$

donde \bar{u}_K es el número medio de defectos por unidad de tipo K , es decir:

$$\bar{u}_K = \frac{\sum_{i=1}^m u_i^K}{m} = \frac{\sum_{i=1}^m c_i^K}{mn} \quad \text{para } K = A, B, C, D.$$

Finalmente, los límites de control para el número de deméritos por unidad son:

$$\begin{aligned} LSC &= \bar{u} + k\hat{\sigma}_U \\ LC &= \bar{u} \\ LIC &= \bar{u} - k\hat{\sigma}_U. \end{aligned}$$

Si, por el contrario, las muestras poseen un tamaño variable n_i , $i = 1, \dots, m$, se toma:

$$\bar{u} = 100\bar{u}_A + 50\bar{u}_B + 10\bar{u}_C + \bar{u}_D$$

con \bar{u}_K el número medio de defectos por unidad de tipo K en las m muestras, es decir,

$$\frac{\sum_{i=1}^m c_i^K}{\sum_{i=1}^m n_i} \quad \text{para } K = A, B, C, D$$

y se establecen límites específicos:

$$\begin{aligned} LSC &= \bar{u} + k\sqrt{\frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{n_i}} \\ LC &= \bar{u} \quad i = 1, \dots, m \\ LIC &= \bar{u} - k\sqrt{\frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{n_i}}. \end{aligned}$$

También se puede emplear el tamaño muestral promedio siguiendo el mismo criterio del gráfico p , de manera que los límites de control aproximados son:

$$\begin{aligned} LSC &= \bar{u} + k\sqrt{\frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{\bar{n}}} \\ LC &= \bar{u} \\ LIC &= \bar{u} - k\sqrt{\frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{\bar{n}}}. \end{aligned}$$

Ejemplo 48 *Se desea establecer un gráfico de control de deméritos mediante el plan Dodge-Torrey para el período base de un modelo de batidoras. Para ello, se estudian 3 meses consecutivos, cuyos resultados se muestran en la Tabla 3.7.*

Entonces, se calcula:

$$\bar{u}_A = \frac{1}{750}, \bar{u}_B = \frac{2}{750}, \bar{u}_C = \frac{18}{750}, \bar{u}_D = \frac{25}{750}$$

Se estudia si se puede utilizar el tamaño muestral promedio:

$$\bar{n} = \frac{750}{3} = 250.$$

Como todos los tamaños muestrales se encuentran en el intervalo $250 \pm 0.25 \cdot 250 = (187.5, 312.5)$, se puede utilizar el tamaño promedio $\bar{n} = 250$.

Entonces:

$$\begin{aligned} \bar{u} &= 100\bar{u}_A + 50\bar{u}_B + 10\bar{u}_C + \bar{u}_D = \frac{405}{750} = 0.54 \\ \hat{\sigma}_u^2 &= \frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{\bar{n}} = \frac{22.43}{250} = 0.0897. \end{aligned}$$

Por tanto, los límites de control 3σ son:

$$\begin{aligned} LSC &= \bar{u} + 3\hat{\sigma}_u = 0.809 \\ LC &= \bar{u} = 0.54 \\ LIC &= \bar{u} - 3\hat{\sigma}_u = 0.271. \end{aligned}$$

Se calcula el valor de u para cada mes:

$$\begin{aligned} u_{Enero} &= 100\frac{0}{250} + 50\frac{0}{250} + 10\frac{6}{250} + \frac{12}{250} = 0.288 \\ u_{Febrero} &= 100\frac{1}{200} + 50\frac{2}{200} + 10\frac{4}{200} + \frac{10}{200} = 1.25 \\ u_{Marzo} &= 100\frac{0}{300} + 50\frac{0}{300} + 10\frac{8}{300} + \frac{3}{300} = 0.276. \end{aligned}$$

Se observa que el mes de febrero se sale de los límites de control, por lo que es conveniente investigar la causa por la que en ese mes se produce un número de defectos fuera de los límites de control.

Una vez conocido que la causa es el cambio de maquinaria, se puede eliminar esa muestra del estudio para obtener los límites de control para el periodo de vigilancia.

Entonces:

$$\bar{u}_A = \frac{0}{750}, \bar{u}_B = \frac{0}{750}, \bar{u}_C = \frac{14}{750}, \bar{u}_D = \frac{15}{750}.$$

Se estudia si se puede utilizar el tamaño muestral promedio:

$$\bar{n} = \frac{550}{2} = 275.$$

Como todos los tamaños muestrales se encuentran en el intervalo $270 \pm 0.25 \cdot 275 = (206.25, 343.75)$, se puede utilizar el tamaño promedio $\bar{n} = 275$ y entonces:

$$\begin{aligned} \bar{u} &= 100\bar{u}_A + 50\bar{u}_B + 10\bar{u}_C + \bar{u}_D = \frac{155}{550} = 0.282 \\ \hat{\sigma}_u^2 &= \frac{100^2\bar{u}_A + 50^2\bar{u}_B + 10^2\bar{u}_C + \bar{u}_D}{\bar{n}} = \frac{1.887}{250} = 0.0075. \end{aligned}$$

Por tanto, los límites de control 3σ son:

$$\begin{aligned} LSC &= \bar{u} + 3\hat{\sigma}_u = 0.3045 \\ LC &= \bar{u} = 0.282 \\ LIC &= \bar{u} - 3\hat{\sigma}_u = 0.2595 \end{aligned}$$

El proceso ya es estable y estos límites se consideran durante el periodo de vigilancia.

3.6 Ejercicios

1. Los datos de la inspección de los interruptores fabricados durante 12 días son:

Día	Inspecciones	Defectuosos
1	1200	4
2	1300	5
3	1250	7
4	1175	6
5	1400	8
6	1350	8
7	1300	5
8	1350	9
9	1400	10
10	1250	7
11	1200	16
12	1150	8

Calcule un gráfico de control para el periodo base del porcentaje de interruptores defectuosos. Si el gráfico no está bajo control, suponga que se han determinado las causas y se han eliminado. Establezca los límites de control para el periodo de vigilancia. Si el porcentaje medio de defectuosos el día siguiente fue del 0.72%, ¿es significativamente diferente del porcentaje estándar con un nivel de significación del 1%?

2. Se estudia el proceso de fabricación de bombillas. Los siguientes datos indican los resultados de las especificaciones de dichas bombillas el

mes de julio de 2003:

Día	% defectuosas	Día	% defectuosas
1	5.5	17	7.4
2	8.3	20	6.3
3	6	21	5
6	5.2	22	5.4
7	4.5	23	7.5
8	6.2	24	7.6
9	5.8	27	5.2
10	7.3	28	6.3
13	4.6	29	5.8
14	6.8	30	6.2
15	7.8	31	6
16	11.6		

Calcule un gráfico de control del porcentaje de bombillas defectuosas.

3. Se estudia el proceso de fabricación de un determinado tipo de ratones atendiendo al turno y al día:

Día	Turno	nº ratones	ratones defectuosos
Lunes	1	138	3
	2	140	5
	3	124	1
Martes	1	137	2
	2	143	6
	3	122	0
Miércoles	12	136	4
	2	148	5
	3	125	1
Jueves	1	138	3
	2	15	3
	3	124	2
Viernes	1	136	4
	2	145	6
	3	123	0

- (a) Calcule los límites de control 3σ .

- (b) Calcule los límites de control para cada turno.
- (c) El lunes siguiente, en el turno 3 se revisan 125 ratones, de los que 3 son defectuosos. Obtenga conclusiones a partir de los resultados de cada uno de los apartados anteriores.
4. Construya un gráfico $100p$ del periodo base y obtenga un valor estándar de p a partir de los siguientes datos de inspección de un determinado producto:

Muestra	Tamaño	Defectuosos
1	500	6
2	600	9
3	600	4
4	300	9
5	300	3
6	600	3
7	200	1
8	500	2
9	200	1
10	200	2

5. Se revisan los coches fabricados de dos modelos distintos buscando defectos en la pintura final. Cada coche se considera un lote independiente:

Nº coche	Modelo	Defectos	Nº coche	Modelo	Defectos
1	Focus	16	11	Ka	9
2	Focus	14	12	Ka	7
3	Focus	19	13	Ka	9
4	Focus	28	14	Ka	11
5	Focus	15	15	Ka	8
6	Focus	22	16	Ka	10
7	Focus	21	17	Ka	11
8	Focus	12	18	Ka	9
9	Focus	23	19	Ka	9
10	Focus	25	20	Ka	7

- (a) Calcule los límites de control de un gráfico c para estos datos.

- (b) ¿Es adecuado tener un único gráfico de control para los dos modelos?
6. Se estudia el número de erratas en el mecanografiado de un documento notarial. La siguiente tabla muestra los resultados de las primeras 24 páginas del documento:

Página	Erratas	Página	Erratas
1	33	13	28
2	27	14	33
3	17	15	36
4	22	16	43
5	1	17	28
6	28	18	30
7	31	19	20
8	38	20	17
9	24	21	23
10	12	22	27
11	41	23	12
12	18	24	15

Compruebe si es estable el gráfico c del proceso. ¿Cuál es el valor estándar de c para el periodo de vigilancia?

7. Una fábrica de productos de fontanería utiliza un gráfico de control para vigilar los defectos en sus productos terminados. Se estudian los datos relativos a 20 días:

Día	productos revisados	defectos	Día	productos revisados	defectos
1	18	12	11	18	8
2	18	14	12	18	14
3	24	20	13	18	9
4	22	18	14	20	10
5	22	15	15	20	14
6	22	12	16	20	13
7	20	11	17	24	16
8	20	15	18	24	18
9	20	12	19	22	20
10	20	10	20	21	17

Obtenga un gráfico de control de disconformidades por producto y el valor estándar para el periodo de vigilancia. Obtenga los límites de aviso.

8. Como ya es conocido, el valor de la constante k está íntimamente relacionado con el nivel de significación mediante la relación $k = z_{1-\frac{\alpha}{2}}$. Con los datos del Ejercicio 7, construya el gráfico de control para un nivel de significación del 5%.
9. Construya un gráfico de control para el periodo base de deméritos por unidad para los circuitos electrónicos fabricados, usando pesos de 50, 10, 5 y 1 para los defectos críticos, importantes, de menor importancia e incidentales, respectivamente.

Mes	Tamaño	críticos	importantes	menor importancia	incidentales
1	4	1	3	2	5
2	4	0	2	4	7
3	4	0	1	3	7
4	4	0	2	1	4
5	4	0	3	5	5
6	4	0	2	3	4
7	4	0	0	4	6
8	4	1	1	1	2
9	4	0	1	3	4
10	4	0	2	2	6

Obtenga un estándar realista para el periodo de vigilancia.

<i>Día</i>	<i>artículos</i>	<i>artículos disconformes</i>	<i>Fracción de disconformes</i>	z_i
1	200	1	0.005	-0.64
2	200	3	0.015	0.82
3	200	0	0	-1.38
4	200	7	0.035	3.75
5	200	2	0.01	0.09
6	200	0	0	-1.38
7	200	1	0.005	-0.64
8	200	0	0	-1.38
9	200	9	0.045	5.22
10	200	3	0.015	0.82
11	200	2	0.01	0.09
12	200	0	0	-1.38
13	200	1	0.005	-0.64
14	200	0	0	-1.38
15	200	4	0.02	1.55
16	200	1	0.005	-0.64
17	200	5	0.025	2.29
18	200	1	0.005	-0.64
19	200	1	0.005	-0.64
20	200	2	0.01	0.09
21	200	0	0	-1.38
22	200	2	0.01	0.09
23	200	1	0.005	-0.64
24	200	1	0.005	-0.64
25	200	0	0	-1.38
<i>Total</i>	5000	47	0.235	

Tabla 3.1: Datos de artículos disconformes en febrero

<i>Día</i>	<i>Tamaño muestral</i>	<i>Nº de disc.</i>	p_i	<i>LIC</i>	<i>LSC</i>	z_i
1	10	1	0.1	0	0.3678	-0.01
2	8	0	0	0	0.4003	-1.05
3	8	1	0.125	0	0.4003	0.26
4	10	0	0	0	0.3678	-1.05
5	11	1	0.091	0	0.3551	-0.10
6	11	2	0.182	0	0.3551	0.85
7	10	1	0.1	0	0.3678	-0.01
8	10	1	0.1	0	0.3678	-0.01
9	9	1	0.111	0	0.3827	0.11
10	9	0	0	0	0.3827	-1.05
11	11	2	0.182	0	0.3551	0.85
12	12	2	0.167	0	0.3439	0.70
13	12	1	0.083	0	0.3439	-0.18
14	12	0	0	0	0.3439	-1.05
15	8	1	0.125	0	0.4003	0.26
16	9	2	0.222	0	0.3827	1.28
17	8	0	0	0	0.4003	-1.05
18	10	1	0.1	0	0.3678	-0.01
19	10	1	0.1	0	0.3678	-0.01
20	11	2	0.182	0	0.3551	0.85
21	8	1	0.125	0	0.4003	0.26
22	9	0	0	0	0.3827	-1.05
23	12	3	0.25	0	0.3439	1.57
24	11	1	0.091	0	0.3551	-0.10
25	10	0	0	0	0.3678	-1.05
<i>Totales</i>	249	25				

Tabla 3.2: Tamaño muestral y número de disconformidades del producto en febrero

Nº muestra	Nº disconf	z_i	Nº muestra	Nº disconf	z_i
1	20	0.32	12	15	-0.84
2	24	1.24	13	8	-2.46
3	13	-1.30	14	19	0.08
4	16	-0.61	15	12	-1.54
5	18	-0.15	16	20	0.32
6	12	-1.54	17	17	-.038
7	10	-2.00	18	30	2.63
8	22	0.78	19	26	1.71
9	27	1.94	20	24	1.24
10	16	-0.61	21	18	-0.15
11	24	1.24	22	19	0.08

Tabla 3.3: Número de disconformidades en muestras de 200 artículos

Nº muestra	Tamaño muestral	Nº disconformidades	Nº disconformidades
	n	c	por unidad, $u = c/n$
1	5	8	1.6
2	5	11	2.2
3	5	12	2.4
4	5	13	2.6
5	5	9	1.8
6	5	12	2.4
7	5	14	2.8
8	5	8	1.6
9	5	11	2.2
10	5	12	2.4
11	5	7	1.4
12	5	10	2
13	5	11	2.2
14	5	14	2.8
15	5	13	2.6
16	5	8	1.6
17	5	9	1.8
18	5	13	2.6
19	5	9	1.8
20	5	7	1.4
Totales	100	211	42.2

Tabla 3.4: Número de disconformidades en lotes de 5 lavadoras

Nº de libro	Nº erratas	Nº capítulos	Nº de erratas por capítulo
1	11	10	1.1
2	16	8	2
3	21	13	1.61
4	14	10	1.4
5	9	9	1
6	12	10	1.2
7	16	12	1.33
8	14	10	1.4
9	23	12	1.92
10	20	12	1.66
Totales	156	106	

Tabla 3.5: Número de erratas y de páginas por capítulo

Nº de libro	n_i	u_i	LIC	LSC	z_i
1	10	1.1	0.34	2.66	-1.03
2	8	2	0.20	2.80	1.15
3	13	1.61	0.48	2.52	0.34
4	10	1.4	0.34	2.66	-0.26
5	9	1	0.27	2.72	-1.22
6	10	1.2	0.34	2.66	-0.77
7	12	1.33	0.44	2.56	-0.47
8	10	1.4	0.34	2.66	-0.25
9	12	1.92	0.44	2.56	1.18
10	12	1.66	0.44	2.56	0.47

Tabla 3.6: Límites de control para el número de erratas y de páginas por capítulo

Tipo defecto	Cantidad			Totales
	Enero	Febrero	Marzo	
<i>A</i>	0	1	0	1
<i>B</i>	0	2	0	2
<i>C</i>	6	4	8	18
<i>D</i>	12	10	3	25
Tamaño	250	200	300	750

Tabla 3.7: Número de defectos de cada tipo de una batidora durante 3 meses

Capítulo 4

Gráficos de Control para Características Variables de Calidad

4.1 Introducción

Aquellas características de calidad que se pueden expresar numéricamente, es decir, que son medibles, se denominan *variables*. Por ejemplo, son características variables de calidad el peso, la temperatura, la presión, la resistencia a la tensión, la dureza, la acidez, etc. En ocasiones, características que normalmente se consideran como atributos (variables cualitativas), pueden ser convertidas en una escala de variables. Consideremos, a modo de ilustración, una gama de rojos; además de calificar a uno de esos rojos como claro u oscuro se pueden asignar valores numéricos con objeto de tener una gradación de la variable.

Los gráficos de control son uno de los métodos más utilizados de control de calidad para características variables. Se ha convertido en un instrumento de gran importancia y sumamente efectivo para controlar la calidad de un determinado proceso durante su desarrollo.

4.2 Gráficos de Control

En la mayoría de los casos, al medir una característica de calidad, ésta se distribuye según una Normal. No obstante, para tamaños suficientemente

elevados, es aplicable el teorema central del límite, con lo que el estadístico media muestral, tendrá, aproximadamente, una distribución Normal. Esta distribución queda completamente determinada conocidos dos parámetros, μ y σ^2 . Por esta causa, se necesitan dos gráficos para controlar una característica variable: uno que permite vigilar la media del proceso y otro para vigilar su variabilidad.

Supóngase que con objeto de determinar si un proceso está o no bajo control estadístico se mide una característica de calidad X con media μ y desviación típica σ . A continuación se distinguen dos casos dependiendo de si dichos parámetros son conocidos o no.

4.2.1 Parámetros Conocidos o Estándares

En ocasiones los parámetros de la característica de calidad son conocidos, bien porque se disponga de estos datos de estudios anteriores, o bien porque vengan dados por la Administración.

Gráfico para la Media

Sea X_1, X_2, \dots, X_n una muestra aleatoria simple de X . El estadístico media muestral \bar{X} se distribuye según una Normal con parámetros $\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Entonces, considérese como estadístico W para establecer los límites de control el propio estadístico media muestral. En consecuencia, los límites de control del gráfico para la media son los siguientes:

$$LSC = \mu_{\bar{X}} + k\sigma_{\bar{X}} = \mu + k\frac{\sigma}{\sqrt{n}}$$

$$LC = \mu_{\bar{X}} = \mu$$

$$LIC = \mu_{\bar{X}} - k\sigma_{\bar{X}} = \mu - k\frac{\sigma}{\sqrt{n}}$$

Si tomamos los límites estándares con $k = 3$ éstos se expresan como

$$\boxed{\mu \pm A\sigma}$$

donde $A = \frac{3}{\sqrt{n}}$.

Nota 49 La constante A sólo depende del tamaño muestral n y sus valores están tabulados (véase Tabla 4.6).

Gráfico para la Variabilidad

Para establecer un gráfico de control para la desviación del proceso σ pueden emplearse dos métodos diferentes:

- a) a partir del recorrido de la muestra, o bien
- b) a partir de la desviación típica muestral.

a) Gráfico R

Sea A la variable aleatoria *amplitud relativa* definida como

$$A = \frac{R}{\sigma},$$

donde R es el recorrido o rango dado por

$$R = X_{(n)} - X_{(1)},$$

con $X_{(1)}$ y $X_{(n)}$ el mínimo y el máximo, respectivamente, de la muestra aleatoria simple X_1, \dots, X_n de X :

$$X_{(1)} = \min_{1 \leq i \leq n} X_i, \quad X_{(n)} = \max_{1 \leq i \leq n} X_i.$$

Las funciones de densidad de los estadísticos máximo y mínimo son, respectivamente:

$$\begin{aligned} f_{X_{(1)}}(x) &= n [1 - F(x)]^{n-1} f(x) \\ f_{X_{(n)}}(x) &= n F^{n-1}(x) f(x) \end{aligned}$$

con $f(x)$ y $F(x)$ las funciones de densidad y de distribución de la variable aleatoria $X \sim \mathcal{N}(\mu, \sigma)$. Entonces

$$E[A] = \frac{E[X_{(n)}] - E[X_{(1)}]}{\sigma} = \frac{n}{\sigma} \left[\int_{\mathbb{R}} x \left\{ F^{n-1}(x) - [1 - F(x)]^{n-1} \right\} f(x) dx \right],$$

integral que se resuelve mediante métodos numéricos. Así, la esperanza de la variable amplitud relativa se denota como d_2 y obsérvese que depende de n (Tabla 4.6). Asimismo

$$Var[A] = \frac{Var[X_{(n)} - X_{(1)}]}{\sigma^2} = \frac{E[(X_{(n)} - X_{(1)})^2] - E[X_{(n)} - X_{(1)}]^2}{\sigma^2}$$

también requiere de métodos numéricos para su resolución. La desviación típica de A se denota d_3 y sus valores están tabulados en función de n (Tabla 4.6).

Nota 50 Las constantes d_2 y d_3 están tabuladas considerando que la distribución de frecuencia es Normal. Sin embargo, Burr (1967) demostró que estas constantes varían muy poco cuando $n \leq 10$, para una amplia variedad de distribuciones.

Considérese que el estadístico objeto del gráfico de control es $W = R$. Se estudian su media y su desviación:

$$\begin{aligned}\mu_R &= \sigma\mu_A = \sigma d_2 \\ \sigma_R &= \sigma\sigma_A = \sigma d_3.\end{aligned}$$

Por tanto, los límites de control para el gráfico de R vienen dados por:

$$\begin{aligned}LSC &= \mu_R + k\sigma_R = \sigma d_2 + k\sigma d_3 = \sigma (d_2 + kd_3) \\ LC &= \mu_R = \sigma d_2 \\ LIC &= \mu_R - k\sigma_R = \sigma d_2 - k\sigma d_3 = \sigma (d_2 - kd_3).\end{aligned}$$

En el caso $k = 3$ se denotan $D_1 = d_2 - 3d_3$ y $D_2 = d_2 + 3d_3$, valores que tan sólo dependen del tamaño muestral n y que aparecen en la Tabla 4.6, de forma que los límites se expresan finalmente como

$\begin{aligned}LSC &= \sigma D_2 \\ LC &= \sigma d_2 \\ LIC &= \sigma D_1.\end{aligned}$

b) Gráfico S

Seguidamente se obtiene un gráfico para controlar la variabilidad del proceso mediante el estadístico desviación típica muestral, S , definido como

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

Es conocido que el estimador máximo verosímil para la varianza del proceso σ^2 viene dado por

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Sin embargo éste no es insesgado, de modo que se introduce un nuevo estimador que sí satisface la propiedad de insesgades:

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Así pues, se establece la siguiente relación entre ambos estimadores

$$(n-1) S^2 = n\tilde{\sigma}^2$$

de donde, dividiendo entre σ^2 , se tiene que

$$\frac{(n-1) S^2}{\sigma^2} = \frac{n\tilde{\sigma}^2}{\sigma^2} \rightsquigarrow \chi^2(n-1).$$

Entonces, la raíz cuadrada positiva de este estadístico se distribuye según una Chi o Normal Positiva¹, esto es

$$\frac{\sqrt{n-1} S}{\sigma} \rightsquigarrow \chi(n-1).$$

En consecuencia,

$$E \left[\frac{\sqrt{n-1} S}{\sigma} \right] = c_2 \sqrt{n} \Rightarrow E[S] = \sigma c_2 \sqrt{\frac{n}{n-1}} = \sigma c_4 \text{ con } c_4 = c_2 \sqrt{\frac{n}{n-1}}$$

$$Var \left[\frac{\sqrt{n-1} S}{\sigma} \right] = n-1 - c_2^2 n \Rightarrow Var[S] = \sigma^2 \left(1 - \frac{c_2^2 n}{n-1} \right) = \sigma^2 (1 - c_4^2).$$

¹Si $X \rightsquigarrow \chi^2(n-1)$ entonces $Y = +\sqrt{X} \rightsquigarrow \chi(n-1)$, denominada distribución Chi o Normal Positiva cuya media y varianza vienen dadas por

$$E[Y] = \sqrt{2} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} = c_2 \sqrt{n}, \quad c_2 = \sqrt{\frac{2}{n}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}$$

$$Var[Y] = n-1 - c_2^2 n$$

Nota 51 Para $n \leq 25$ los valores de c_4 están tabulados (Tabla 4.6). Si $n > 25$ esta cantidad se aproxima por

$$c_4 \cong \frac{4(n-1)}{4n-3}. \quad (4.1)$$

Una vez estudiadas algunas características del estadístico S se establecen los límites de control para el gráfico S tomando, por supuesto, como estadístico de interés el propio S , es decir, $W = S$:

$$LSC = \mu_S + k\sigma_S = \sigma c_4 + k\sigma\sqrt{1-c_4^2} = \sigma \left(c_4 + k\sqrt{1-c_4^2} \right)$$

$$LC = \mu_S = \sigma c_4$$

$$LIC = \mu_S - k\sigma_S = \sigma c_4 - k\sigma\sqrt{1-c_4^2} = \sigma \left(c_4 - k\sqrt{1-c_4^2} \right).$$

Particularizando al caso más habitual $k = 3$ y llamando

$$B_5 = c_4 - 3\sqrt{1-c_4^2} \text{ y } B_6 = c_4 + 3\sqrt{1-c_4^2},$$

los límites para el gráfico de S quedan

$\begin{aligned} LSC &= \sigma B_6 \\ LC &= \sigma c_4 \\ LIC &= \sigma B_5. \end{aligned}$

Nota 52 Los valores de las constantes B_5 y B_6 para cada $n \leq 25$ también aparecen en la Tabla 4.6. Si $n > 25$ se calculan de forma aproximada teniendo en cuenta (4.1) y tomando

$$B_5 \cong c_4 - \frac{3}{\sqrt{2(n-1)}} \text{ y } B_6 \cong c_4 + \frac{3}{\sqrt{2(n-1)}}.$$

¿Gráfico R o Gráfico S ? Cabe plantearse la pregunta de qué gráfico será más eficiente para controlar la variabilidad de un proceso, si el gráfico de R o el de S . Si el tamaño muestral es relativamente pequeño, el rango o recorrido resulta un estimador eficaz y sencillo para la desviación típica, en cuyo caso el método del recorrido produce un estimador casi tan bueno como el estimador cuadrático habitual. En el caso de valores moderados de n ($n \geq 10$), el recorrido pierde rápidamente su eficiencia, al no tener en cuenta toda la información de la muestra, de modo que será conveniente representar el gráfico de S .

4.2.2 Parámetros Desconocidos

El caso más habitual, sin embargo, es que los valores de los parámetros del proceso o de la característica de calidad en cuestión sean desconocidos. Por consiguiente el primer paso hacia la construcción de los gráficos de control será la estimación de dichos parámetros. Para ello se extraen m muestras de tamaño n cada una, es decir,

$$\begin{aligned} &x_{11}, x_{12}, \dots, x_{1n} \\ &x_{21}, x_{22}, \dots, x_{2n} \\ &\dots\dots\dots \\ &x_{m1}, x_{m2}, \dots, x_{mn} \end{aligned}$$

Nota 53 *Obsérvese que el primer subíndice hace referencia a la muestra y el segundo al lugar que ocupa la observación dentro de la muestra a la que pertenece.*

Así, la media muestral de la muestra i -ésima es

$$\bar{x}_i = \frac{x_{i1} + x_{i2} + \dots + x_{in}}{n}, \quad i = 1, \dots, m.$$

Entonces, el mejor estimador de la media μ viene dado por

$$\hat{\mu} = \bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \dots + \bar{x}_m}{m}.$$

Al igual que en el caso de parámetros conocidos, la desviación típica puede estimarse a partir de

- las amplitudes o recorridos de las m muestras o
- las desviaciones típicas de las m muestras.

Dependiendo de la estimación utilizada se tienen los gráficos de control de \bar{X} y R y de \bar{X} y S , respectivamente, y que se detallan a continuación.

Gráficos de \bar{X} y R

Para obtener los gráficos de \bar{X} y R se requiere un estimador para σ . Ya se vio que $\mu_R = \sigma d_2$, de donde despejando σ se tiene

$$\sigma = \frac{\mu_R}{d_2},$$

lo que proporciona una estimación sencilla para la variabilidad

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

con

$$\bar{R} = \frac{R_1 + R_2 + \dots + R_m}{m},$$

$$R_i = \max_{1 \leq j \leq n} x_{ij} - \min_{1 \leq j \leq n} x_{ij}, \quad i = 1, \dots, m$$

el recorrido o rango de cada muestra y d_2 la media de la variable aleatoria amplitud relativa.

Por tanto, los límites de control para el gráfico de \bar{X} son:

$$\begin{aligned} LSC &= \hat{\mu} + k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} + k \frac{\bar{R}}{d_2 \sqrt{n}} \\ LC &= \hat{\mu} = \bar{\bar{x}} \\ LIC &= \hat{\mu} - k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} - k \frac{\bar{R}}{d_2 \sqrt{n}}. \end{aligned}$$

Si $k = 3$ los límites se expresan de una forma especial. Así, teniendo en cuenta que $A_2 = \frac{3}{d_2 \sqrt{n}}$, se tiene

$$\boxed{\bar{\bar{x}} \pm A_2 \bar{R}.$$

Nota 54 Los valores de A_2 sólo dependen del tamaño muestral n y también están tabulados (Tabla 4.6).

En cuanto al gráfico para la variabilidad, basta con sustituir los valores estándares que aparecen en el gráfico de R por sus estimaciones, esto es

$$\begin{aligned}LSC &= \hat{\sigma} (d_2 + kd_3) = \bar{R} \left(1 + k \frac{d_3}{d_2} \right) \\LC &= \hat{\sigma} d_2 = \bar{R} \\LIC &= \hat{\sigma} (d_2 - kd_3) = \bar{R} \left(1 - k \frac{d_3}{d_2} \right).\end{aligned}$$

De nuevo, si $k = 3$ los límites quedan finalmente como se indica:

$\begin{aligned}LSC &= \bar{R}D_4 \\LC &= \bar{R} \\LIC &= \bar{R}D_3\end{aligned}$

donde $D_3 = 1 - 3 \frac{d_3}{d_2}$ y $D_4 = 1 + 3 \frac{d_3}{d_2}$ están tabulados en función del valor de n (Tabla 4.6).

Nota 55 Las constantes necesarias para establecer límites probabilísticos en el gráfico R también están tabuladas. Así por ejemplo, en ? encontramos tablas para $2 \leq n \leq 12$.

Ejemplo 56 Una compañía de seguros extrae muestras de tamaño 5 cada una sobre las reclamaciones contra la empresa por daños en los automóviles. En la Tabla 4.1 se muestran los valores de la característica de calidad expresados en decenas de euros. Representétese el diagrama de control de \bar{X} y R .

Se calculan la media y recorrido muestrales:

$$\begin{aligned}\hat{\mu} = \bar{\bar{x}} &= \frac{1}{15} \sum_{i=1}^{15} \bar{x}_i = \frac{1221.4}{15} = 81.4267 \\ \bar{R} &= \frac{1}{15} \sum_{i=1}^{15} R_i = \frac{1101}{15} = 73.4.\end{aligned}$$

Entonces, los límites de control para el gráfico R son:

$$\begin{aligned}LSC &= \bar{R}D_4 = 73.4 \cdot 2.114 = 155.1676 \\LC &= \bar{R} = 73.4 \\LIC &= \bar{R}D_3 = 0.\end{aligned}$$

N° de muestra	Valor de la reclamación					\bar{X}	R
	1	2	3	4	5		
1	49	47	47	96	69	61.6	49
2	93	124	59	102	74	90.4	65
3	43	69	73	80	78	68.6	37
4	101	45	84	28	21	55.8	80
5	101	81	55	44	98	75.8	57
6	119	45	77	102	80	84.6	74
7	73	103	104	98	48	85.2	56
8	71	80	120	83	119	94.6	49
9	15	71	200	77	77	88.0	185
10	61	78	47	80	82	69.6	35
11	58	47	105	73	91	74.8	58
12	54	99	64	109	73	79.8	55
13	46	78	87	112	59	76.4	66
14	225	120	50	153	76	124.8	175
15	98	119	104	59	77	91.4	60

Tabla 4.1: Medias y recorridos muestrales para las reclamaciones en una empresa aseguradora de automóviles

Y el gráfico queda como se aprecia en la Figura 4.1.

Puede observarse que los recorridos correspondientes a las muestras número 9 y 14 revelan que el proceso está fuera de control, ya que ambas se sitúan por encima del límite superior, lo cual puede evidenciar la presencia de una causa atribuible. Si se construye el gráfico para la media (Figura 4.2), se aprecia que es la muestra 14 la que, en este caso, da una condición fuera de control:

$$LSC = \bar{\bar{x}} + A_2\bar{R} = 81.4267 + 0.577 \cdot 73.4 = 123.7785$$

$$LC = \bar{\bar{x}} = 81.4267$$

$$LIC = \bar{\bar{x}} - A_2\bar{R} = 81.4267 - 0.577 \cdot 73.4 = 39.0749.$$

En consecuencia, tampoco la media del proceso es estable a lo largo del tiempo, es decir, el proceso, en su conjunto, no está bajo control estadístico.

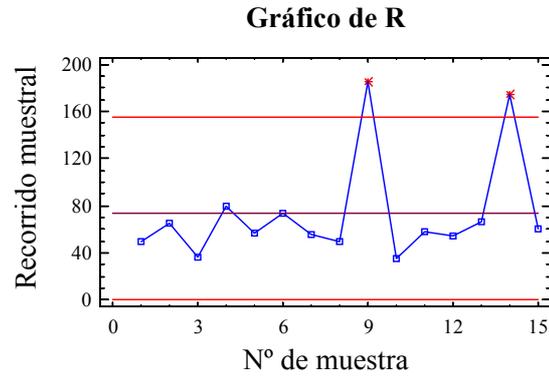


Figura 4.1: Gráfico de R para las reclamaciones en una empresa aseguradora de automóviles

Gráficos de \bar{X} y S

En primer lugar se obtiene un estimador para σ a partir de las desviaciones típicas muestrales. En el caso de parámetros conocidos se sabe que $\mu_S = \sigma c_4$; de aquí utilizando estimaciones en ambos miembros, se tiene un nuevo estimador para la variabilidad del proceso:

$$\hat{\mu}_S = \hat{\sigma} c_4 \Rightarrow \hat{\sigma} = \frac{\hat{\mu}_S}{c_4} = \frac{\bar{S}}{c_4}$$

donde

$$\bar{S} = \frac{1}{m} \sum_{i=1}^m S_i$$

y S_i es la desviación típica muestral asociada a la muestra i -ésima

$$S_i = \sqrt{\frac{j=1 \sum^n (x_{ij} - \bar{x}_i)^2}{n-1}}, \quad i = 1, \dots, m.$$

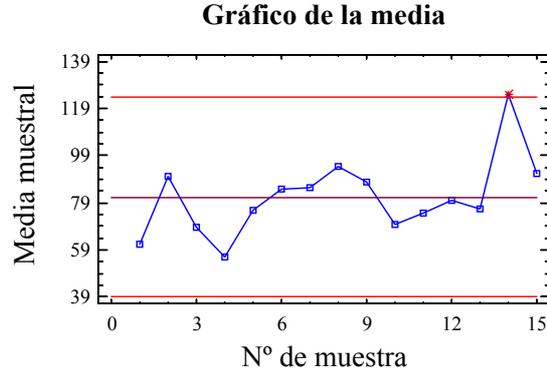


Figura 4.2: Gráfico de \bar{X} para las reclamaciones en una empresa aseguradora de automóviles

Entonces, los límites de control para el gráfico de \bar{X} vienen dados por:

$$LSC = \hat{\mu} + k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} + k \frac{\bar{S}}{c_4 \sqrt{n}} = \bar{\bar{x}} + \bar{S} \frac{k}{c_4 \sqrt{n}}$$

$$LC = \hat{\mu} = \bar{\bar{x}}$$

$$LIC = \hat{\mu} - k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} - k \frac{\bar{S}}{c_4 \sqrt{n}} = \bar{\bar{x}} - \bar{S} \frac{k}{c_4 \sqrt{n}}$$

En el caso particular en el que $k = 3$, se denomina A_3 a $\frac{3}{c_4 \sqrt{n}}$, de modo que los límites de control anteriores pueden expresarse abreviadamente como:

$$\boxed{\bar{\bar{x}} \pm A_3 \bar{S}}$$

Para controlar la variabilidad, se retoman los límites para el gráfico de S con parámetros conocidos y éstos se sustituyen por sus correspondientes

estimaciones, esto es

$$LSC = \hat{\sigma} \left(c_4 + k\sqrt{1 - c_4^2} \right) = \frac{\bar{S}}{c_4} \left(c_4 + k\sqrt{1 - c_4^2} \right) = \bar{S} \left(1 + k\frac{\sqrt{1 - c_4^2}}{c_4} \right)$$

$$LC = \hat{\sigma}c_4 = \bar{S}$$

$$LIC = \hat{\sigma} \left(c_4 - k\sqrt{1 - c_4^2} \right) = \frac{\bar{S}}{c_4} \left(c_4 - k\sqrt{1 - c_4^2} \right) = \bar{S} \left(1 - k\frac{\sqrt{1 - c_4^2}}{c_4} \right).$$

Sin más que calcular los límites de 3σ ($k = 3$) y notando

$$B_3 = 1 - 3\frac{\sqrt{1 - c_4^2}}{c_4} \quad \text{y} \quad B_4 = 1 + 3\frac{\sqrt{1 - c_4^2}}{c_4},$$

los límites de control para el gráfico de S se reescriben como

$\begin{aligned} LSC &= \bar{S}B_4 \\ LC &= \bar{S} \\ LIC &= \bar{S}B_3. \end{aligned}$
--

Nota 57 La Tabla 4.6 contiene los valores de estas constantes para valores de $n \leq 25$. Si $n > 25$ se han de calcular de forma aproximada a partir de (4.1) y considerando

$$B_3 \cong 1 - \frac{3}{c_4\sqrt{2(n-1)}} \quad \text{y} \quad B_4 \cong 1 + \frac{3}{c_4\sqrt{2(n-1)}}.$$

Un resumen de los métodos de cálculo de los límites de control para los diagramas \bar{X} , R y S aparece en la Tabla 4.2.

Ejemplo 58 Con los datos del ejemplo anterior, se desea representar los gráficos para la media y la desviación típica a partir de las desviaciones típicas muestrales. En este caso, la estimación de la media no varía, pero sí la de la variabilidad, para lo cual es necesario calcular, previamente, todas las desviaciones típicas muestrales (véase Tabla 4.3):

$$\bar{\bar{x}} = 81.4267$$

$$\bar{S} = \frac{1}{15} \sum_{i=1}^{15} S_i = \frac{445.201}{15} = 29.6801.$$

μ y σ conocidos			
Diagrama	LSC	LC	LIC
\bar{X}	$\mu + A\sigma$	μ	$\mu - A\sigma$
R	$D_2\sigma$	$d_2\sigma$	$D_1\sigma$
S	$B_6\sigma$	$c_4\sigma$	$B_5\sigma$

μ y σ desconocidos			
Diagrama	LSC	LC	LIC
\bar{X} y R $\left\{ \begin{array}{l} \bar{X} \\ R \end{array} \right.$	$\bar{\bar{x}} + A_2\bar{R}$ $D_4\bar{R}$	$\bar{\bar{x}}$ \bar{R}	$\bar{\bar{x}} - A_2\bar{R}$ $D_3\bar{R}$
\bar{X} y S $\left\{ \begin{array}{l} \bar{X} \\ S \end{array} \right.$	$\bar{\bar{x}} + A_3\bar{S}$ $B_4\bar{S}$	$\bar{\bar{x}}$ \bar{S}	$\bar{\bar{x}} - A_3\bar{S}$ $B_3\bar{S}$

Tabla 4.2: Límites para los diagramas \bar{X} , R y S

Los límites de control para el gráfico de S son:

$$LSC = \bar{S}B_4 = 29.6801 \cdot 2.089 = 62.0017$$

$$LC = \bar{S} = 29.6801$$

$$LIC = \bar{S}B_3 = 29.6801 \cdot 0 = 0.$$

La Figura 4.3 muestra el gráfico de S media asociado.

donde se aprecian condiciones fuera de control en las muestras 9 y 14, lo que indica que la variabilidad del proceso no es estable. Análogamente, los límites para el gráfico de la media son:

$$LSC = \bar{\bar{x}} + A_3\bar{S} = 81.4267 + 1.427 \cdot 29.6801 = 123.7802$$

$$LC = \bar{\bar{x}} = 81.4267$$

$$LIC = \bar{\bar{x}} - A_3\bar{S} = 81.4267 - 1.427 \cdot 29.6801 = 39.0732.$$

Gráficamente, se tiene la Figura 4.4.

La muestra 14 se sitúa por encima del límite superior de control, por tanto, la media tampoco es estable y, en consecuencia, el proceso está fuera de control.

Nº de muestra	Reclamación	
	\bar{X}	S
1	61.6	21.3495
2	90.4	25.1257
3	68.6	14.9432
4	55.8	35.1383
5	75.8	25.4892
6	84.6	27.9875
7	85.2	24.3249
8	94.6	23.1582
9	88.0	67.8307
10	69.6	15.1427
11	74.8	23.6263
12	79.8	23.3602
13	76.4	25.5402
14	124.8	68.62
15	91.4	23.5648

Tabla 4.3: Medias y desviaciones típicas muestrales para las reclamaciones de una empresa aseguradora de automóviles

4.3 Interpretación de los Gráficos para la Media y la Variabilidad

Un gráfico de control puede indicar una condición fuera de control sin que un solo punto esté fuera de los límites (basta con que el patrón de los puntos ubicados presente un comportamiento sistemático, como ya se comentó en el Capítulo 2).

Para interpretar patrones de comportamiento en el diagrama de la media, hay que determinar primero si el diagrama para controlar la variabilidad está bajo control o no.

Si ambos diagramas presentan un patrón no aleatorio, la mejor estrategia es eliminar primero las causas atribuibles en el diagrama de la variabilidad. En muchos casos, esto eliminará automáticamente el patrón no aleatorio en el diagrama de la media. Nunca hay que tratar de interpretar el gráfico de la media cuando el de la dispersión indica una condición fuera de control. El procedimiento se resume en la Figura 4.5.

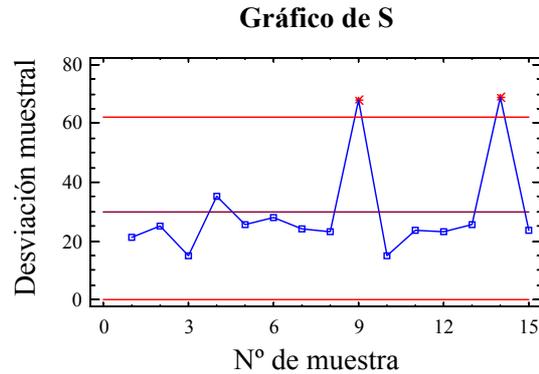


Figura 4.3: Gráfico de S para las reclamaciones en una empresa aseguradora de automóviles

Ejemplo 59 *El Departamento de Control de Calidad de una fábrica de champús desea estudiar el pH de sus productos. Para ello toma 32 muestras de 4 productos cada una y mide el pH. Se obtuvieron los resultados de la Tabla 4.4 ($X = pH \times 10$).*

Paso 1: *Se dispone de 32 muestras para el período base, de modo que se avanza hasta el siguiente paso.*

Paso 2: *En la Tabla 4.5 se calculan inicialmente los valores \bar{X} y R para establecer los límites de control durante el período base.*

$$\bar{\bar{x}} = \frac{1}{32} \sum_{i=1}^{32} \bar{x}_i = \frac{1529,75}{32} = 47.8047$$

$$\bar{R} = \frac{1}{32} \sum_{i=1}^{32} R_i = \frac{382}{32} = 11.9375.$$

Cada muestra es de tamaño 4, de modo que los valores asociados para $k = 3$ son

$$A_2 = 0.729, D_3 = 0 \quad y \quad D_4 = 2.282.$$

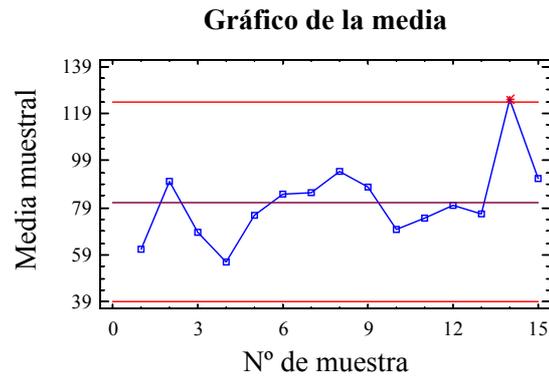


Figura 4.4: Gráfico de \bar{X} para las reclamaciones en una empresa aseguradora de automóviles

Así es posible establecer los límites de control para \bar{X} y R :

$$LSC(R) = D_4\bar{R} = 2.282 \cdot 11.9375 = 27.2435$$

$$LIC(R) = D_3\bar{R} = 0$$

$$LSC(\bar{X}) = \bar{\bar{x}} + A_2\bar{R} = 47.8047 + 0.729 \cdot 11.9375 = 56.5013$$

$$LIC(\bar{X}) = \bar{\bar{x}} - A_2\bar{R} = 47.8047 - 0.729 \cdot 11.9375 = 39.1081.$$

De esta forma queda completado el Paso 2.

Paso 3: A continuación se comprueba si todos los recorridos están dentro de los límites de control establecidos para el gráfico R . En la Figura 4.6 se aprecia cómo la muestra número 13 se sale de estos límites, lo que significa que el proceso no es estable durante el período base. Por tanto, se procede a la eliminación del conjunto de datos de la muestra 13 y se regresa al Paso 2:

Ahora los estadísticos toman los valores

$$\bar{\bar{x}} = \frac{1}{31} \sum_{i=1, i \neq 13}^{32} \bar{x}_i = 47.8145$$

$$\bar{R} = \frac{1}{31} \sum_{i=1, i \neq 13}^{32} R_i = 11.3226$$

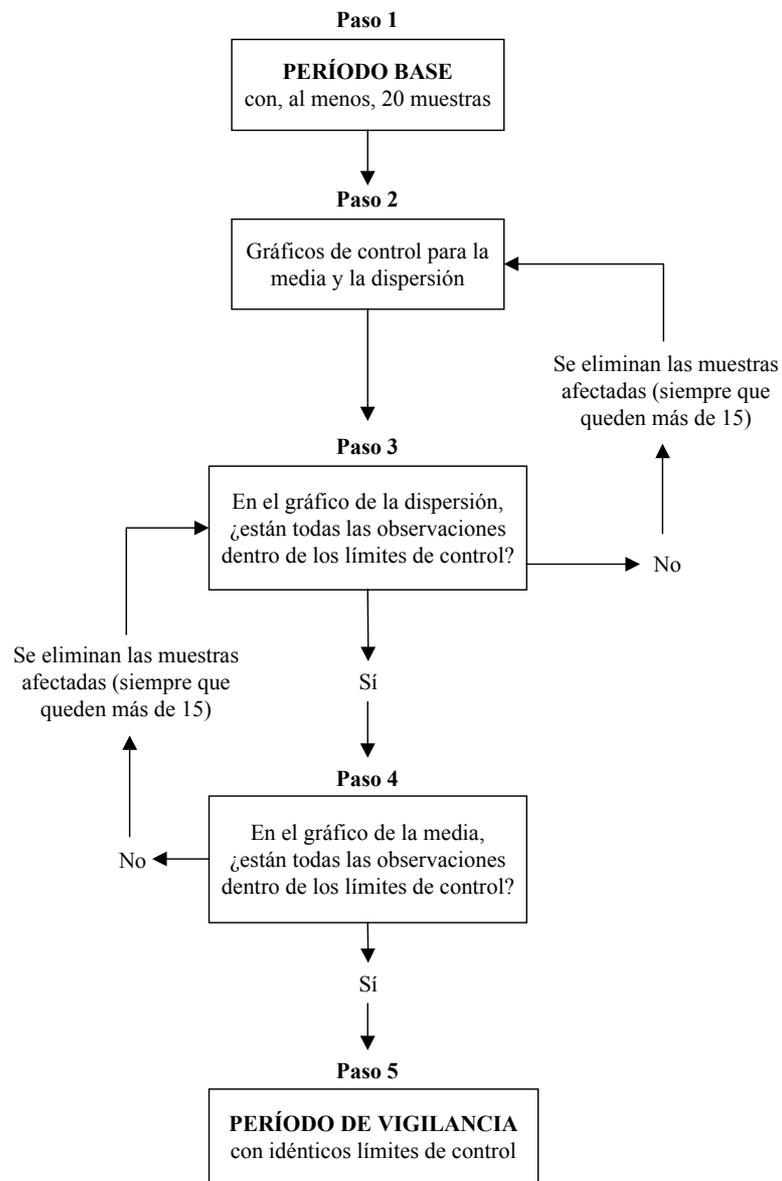


Figura 4.5: Pasos para establecer los límites del período de vigilancia

Nº de muestra	1	2	3	4	Nº de muestra	1	2	3	4
1	42	40	46	55	17	64	66	59	64
2	50	43	32	51	18	51	53	48	41
3	46	44	58	46	19	48	47	50	49
4	59	52	48	45	20	47	46	50	52
5	37	34	43	49	21	50	53	34	55
6	55	58	60	56	22	42	45	50	48
7	47	51	49	40	23	54	40	55	47
8	47	38	45	44	24	57	49	51	50
9	49	40	40	53	25	50	49	50	58
10	54	40	56	51	26	44	46	50	52
11	54	48	51	48	27	40	57	43	36
12	44	51	36	47	28	41	42	40	45
13	68	42	43	37	29	50	44	47	49
14	40	45	45	52	30	41	42	32	30
15	62	41	44	45	31	38	40	39	41
16	56	45	55	45	32	65	54	55	57

Tabla 4.4: PH de los productos de una fábrica de champús

y los nuevos límites de control son²:

$$LSC(R) = D_4\bar{R} = 25.8401$$

$$LIC(R) = D_3\bar{R} = 0$$

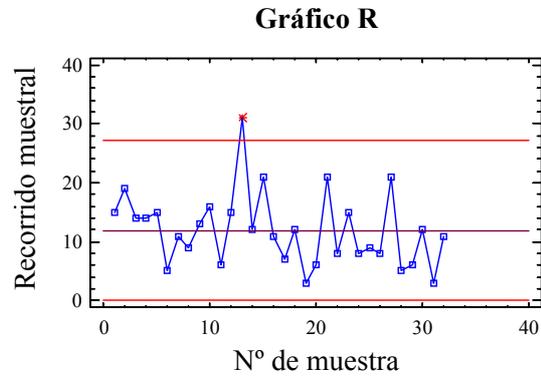
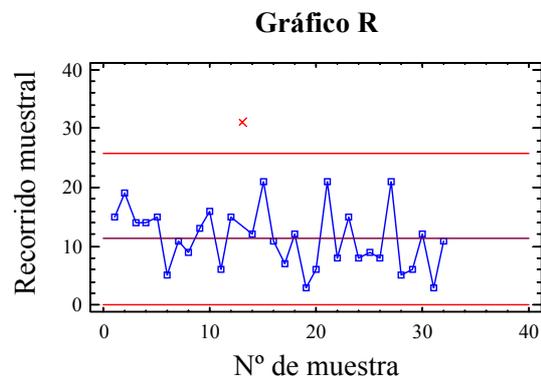
$$LSC(\bar{X}) = \bar{\bar{X}} + A_2\bar{R} = 56.0631$$

$$LIC(\bar{X}) = \bar{\bar{X}} - A_2\bar{R} = 39.5659.$$

De nuevo se comprueba - de acuerdo con el Paso 3 - si todos los recorridos caen dentro de los límites de control. Como la respuesta a esta cuestión es afirmativa (véase la Figura 4.7), se avanza al Paso 4.

Paso 4: Una vez aquí, se observa que las muestras 6, 17, 30, 31 y 32 se encuentran fuera de los límites de control establecidos para \bar{X} (Figura 4.8). En consecuencia, se eliminan estas muestras - quedan, pues, 26 > 15 - y

²Los valores de A_2 , A_3 y D_4 no varían, pues cada muestra sigue teniendo el mismo tamaño (4 productos).

Figura 4.6: Gráfico *R* para los datos sobre el ph de los champúsFigura 4.7: Gráfico *R* una vez eliminada la muestra número 13

Nº de muestra	\bar{x}_i	R_i	Nº de muestra	\bar{x}_i	R_i
1	45.75	15	17	63.25	7
2	44.0	19	18	48.25	12
3	48.5	14	19	48.5	3
4	51.0	14	20	48.75	6
5	40.75	15	21	48.0	21
6	57.25	5	22	46.25	8
7	46.75	11	23	49.0	15
8	43.5	9	24	51.75	8
9	45.5	13	25	51.75	9
10	50.25	16	26	48.0	8
11	50.25	6	27	44.0	21
12	44.5	15	28	42.0	5
13	47.5	31	29	47.5	6
14	45.5	12	30	36.25	12
15	48.0	21	31	39.5	3
16	50.25	11	32	57.75	11

Tabla 4.5: Medias y recorridos para los datos sobre el ph de los champús

se vuelve al Paso 2. Los cálculos son los siguientes:

$$\bar{\bar{x}} = \frac{1}{26} \sum_{i \neq 6, 13, 17}^{i=1}^{29} \bar{x}_i = 47.2404$$

$$\bar{R} = \frac{1}{26} \sum_{i \neq 6, 13, 17}^{i=1}^{29} R_i = 12.0385.$$

Los límites de control para los gráficos \bar{X} y R son:

$$LSC(R) = D_4 \bar{R} = 27.4739$$

$$LIC(R) = D_3 \bar{R} = 0$$

$$LSC(\bar{X}) = \bar{\bar{x}} + A_2 \bar{R} = 56.0105$$

$$LIC(\bar{X}) = \bar{\bar{x}} - A_2 \bar{R} = 38.4703.$$

De acuerdo con los Pasos 3 y 4, todas las observaciones están dentro de los límites de control, de modo que con las 26 muestras finales, el proceso - en el período base - está bajo control.

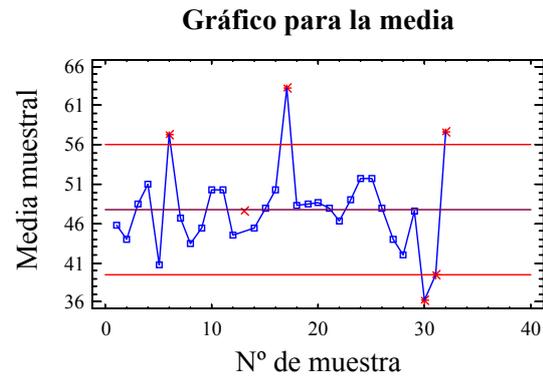


Figura 4.8: Gráfico de \bar{X} una vez eliminada la muestra número 13

En conclusión, se aplican los límites 56.0105, 48 y 38.4703 para \bar{X} y 27.4739 y 0 para R , al período de vigilancia (Figuras 4.9 y 4.10).

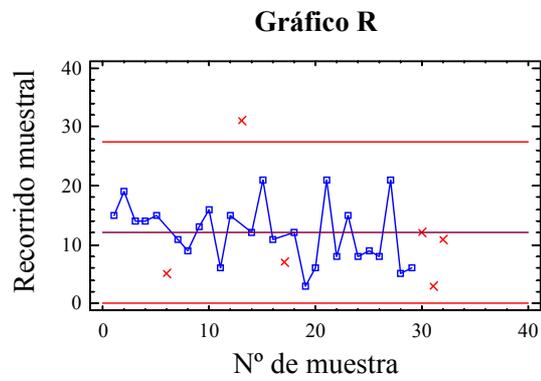


Figura 4.9: Límites de control para la variabilidad del proceso durante el período de vigilancia

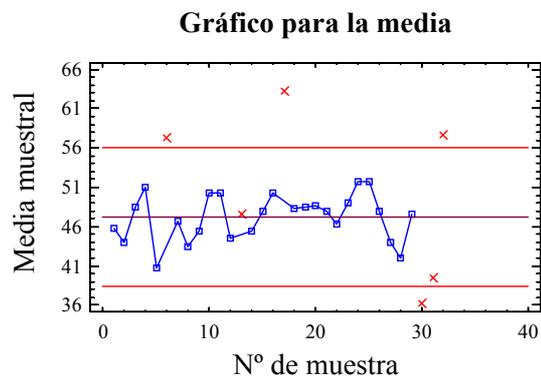


Figura 4.10: Límites de control para la media del proceso durante el período de vigilancia

Tamaño de muestra, n	Gráfico \bar{X}			Gráfico S						Gráfico R					
	A	A_2	A_3	c_4	B_3	B_4	B_5	B_6	d_2	d_3	D_1	D_2	D_3	D_4	
2	2.121	1.880	2.659	0.7979	0	3.267	0	2.606	1.128	0.853	0	3.686	0	3.267	
3	1.732	1.023	1.954	0.8862	0	2.568	0	2.279	1.693	0.888	0	4.358	0	2.574	
4	1.500	0.729	1.628	0.9213	0	2.266	0	2.088	2.059	0.880	0	4.698	0	2.282	
5	1.342	0.577	1.427	0.9400	0	2.089	0	1.964	2.326	0.864	0	4.918	0	2.114	
6	1.225	0.483	1.287	0.9515	0.030	1.970	0.029	1.874	2.534	0.848	0	5.078	0	2.004	
7	1.134	0.419	0.182	0.9594	0.118	1.882	0.113	1.806	2.704	0.833	0.204	5.204	0.076	1.924	
8	1.061	0.373	1.099	0.9650	0.185	1.815	0.179	1.751	2.847	0.820	0.388	5.306	0.136	1.864	
9	1.000	0.337	1.032	0.9693	0.239	1.761	0.232	1.707	2.970	0.808	0.547	5.393	0.184	1.816	
10	0.949	0.308	0.975	0.9727	0.284	1.716	0.276	1.669	3.078	0.797	0.687	5.469	0.223	1.777	
11	0.905	0.285	0.927	0.9754	0.321	1.679	0.313	1.637	3.173	0.787	0.811	5.535	0.256	1.744	
12	0.866	0.266	0.886	0.9776	0.354	1.646	0.346	1.610	3.258	0.778	0.922	5.594	0.283	1.717	
13	0.832	0.249	0.850	0.9794	0.382	1.618	0.374	1.585	3.336	0.770	1.025	5.647	0.307	1.693	
14	0.802	0.235	0.817	0.9810	0.406	1.594	0.399	1.563	3.407	0.763	1.118	5.696	0.328	1.672	
15	0.775	0.223	0.789	0.9823	0.428	1.572	0.421	1.544	3.472	0.756	1.203	5.741	0.347	1.653	
16	0.750	0.212	0.763	0.9835	0.448	1.552	0.440	1.526	3.532	0.750	1.282	5.782	0.363	1.637	
17	0.728	0.203	0.739	0.9845	0.466	1.534	0.458	1.511	3.588	0.744	1.356	5.820	0.378	1.622	
18	0.707	0.194	0.718	0.9854	0.482	1.518	0.475	1.496	3.640	0.739	1.424	5.856	0.391	1.608	
19	0.688	0.187	0.698	0.9832	0.497	1.503	0.490	1.483	3.689	0.734	1.487	5.891	0.403	1.597	
20	0.671	0.180	0.680	0.9869	0.510	1.490	0.504	1.470	3.735	0.729	1.549	5.921	0.415	1.585	
21	0.655	0.173	0.663	0.9876	0.523	1.477	0.516	1.459	3.778	0.724	1.605	5.951	0.425	1.575	
22	0.640	0.167	0.647	0.9882	0.534	1.466	0.528	1.448	3.819	0.720	1.659	5.979	0.434	1.566	
23	0.626	0.162	0.633	0.9887	0.545	1.455	0.539	1.438	3.858	0.716	1.710	6.006	0.443	1.557	
24	0.612	0.157	0.619	0.9892	0.555	1.445	0.549	1.429	3.895	0.712	1.759	6.031	0.451	1.548	
25	0.600	0.153	0.606	0.9896	0.565	1.435	0.559	1.420	3.931	0.708	1.806	6.056	0.459	1.541	

Tabla 4.6: Constantes para construir los gráficos de control de variables con $k = 3$

4.4 Ejercicios

1. Con la finalidad de controlar un proceso con valores estándares $\mu = 15$ y $\sigma = 2$, se extrajeron muestras cuyo tamaño fue de 10 unidades. Entonces:
 - (a) Establezca la línea central y los límites de control para el gráfico.
 - (b) Halle la línea central y los límites de control para el gráfico R .
 - (c) Obtenga la línea central y los límites de control para el gráfico S .

2. A finales de 2003, una gran empresa electrónica europea anunció su intención de competir directamente con la principal del ramo en Estados Unidos, en la fabricación de ciertos componentes electrónicos de precisión. El director ejecutivo de la empresa europea exigió controles exhaustivos de la calidad con objeto de medir la variación en el grosor de uno de sus componentes. El departamento de control de calidad extrajo cada hora muestras de tamaño 5 durante 5 horas. Los resultados (en cm) aparecen en la siguiente tabla:

Hora	1	2	3	4	5
6:00	4.9	4.8	4.5	5.1	6.4
7:00	6.8	5.2	5.1	7.0	5.4
8:00	7.0	6.5	6.0	6.1	6.9
9:00	6.8	6.7	6.2	6.9	6.6
10:00	5.9	4.7	4.4	4.3	5.1

El director ejecutivo ha dado la orden de detener el proceso de producción si se identifica una causa atribuible, hasta que ésta sea corregida. ¿Qué cree que ocurrirá?

3. El Ministerio de Economía y Hacienda estableció que los tiempos de realización de la declaración del IRPF deberían ser estándar, oscilando entre los 10 y los 15 minutos. Se seleccionaron 10 oficinas de la Agencia Tributaria y se midió el tiempo de atención de 5 clientes. Con base en estos resultados, construya e interprete los gráficos de control apropiados para determinar si los tiempos de atención están bajo control. En caso necesario, establezca los límites para el período

de vigilancia. ¿Qué porcentaje de clientes son atendidos dentro de los límites marcados por el Ministerio?

N° de cliente				
1	2	3	4	5
4.1	10.5	9.8	7.9	16.0
3.2	15.0	11.9	9.5	8.9
3.8	9.0	11.1	7.1	7.9
5.5	8.0	6.0	6.0	6.5
5.9	8.7	12.1	7.1	6.6
6.5	4.0	14.2	16.5	6.0
6.7	7.3	10.8	12.4	4.0
9.5	6.1	15.6	14.4	5.2
8.7	7.4	13.0	12.3	6.4

4. Un conocido centro hospitalario de la capital desea llevar a cabo un proyecto para mejorar la calidad de tiempo previo necesario para la admisión de un paciente, para lo cual pretende utilizar gráficos de control. Los datos más recientemente obtenidos son los siguientes:

N° de subgrupo	Observación			N° de subgrupo	Observación		
	X_1	X_2	X_2		X_1	X_2	X_3
1	6.0	5.8	6.1	13	8.4	6.9	7.4
2	5.2	6.4	6.9	14	6.2	5.2	6.8
3	5.5	5.8	5.2	15	4.9	6.6	6.6
4	5.0	5.7	6.5	16	7.0	6.4	6.1
5	6.7	6.5	5.5	17	5.4	6.5	6.7
6	5.8	5.2	5.0	18	6.6	7.0	6.8
7	5.6	5.1	5.2	19	3.8	6.2	7.1
8	6.0	5.8	6.0	20	6.7	5.4	6.7
9	5.5	4.9	5.7	21	6.8	6.5	5.2
10	4.3	6.4	6.3	22	5.9	6.4	6.0
11	6.2	6.9	5.0	23	6.7	6.3	4.6
12	6.7	7.1	6.2	24	7.4	6.8	6.3

- (a) Defina la característica variable de calidad y construir gráficos de control *apropiados*. Razone la respuesta.
- (b) Establezca los límites de control para el período de vigilancia. ¿Cuál es la distribución aproximada del tiempo promedio previo necesario para la admisión de un paciente?

5. En un proceso de envasado de cacao, se extrajeron 40 muestras con una frecuencia regular con objeto de controlar el contenido de cada recipiente. Si se supone que la característica de calidad “peso del recipiente” (en gramos) se distribuye según una normal, se pide:

- (a) Establezca límites de control apropiados para la media y la variabilidad del proceso de envasado si se obtuvieron los siguientes datos:

$$\sum_{i=1}^{40} \bar{x}_i = 9950, \sum_{i=1}^{40} R_i = 20$$

- (b) Supuesto que el gráfico no mostró ninguna situación fuera de control, ¿cuáles son los parámetros estándares del proceso?
- (c) De cara a satisfacer al cliente, se considera insatisfactorio un recipiente cuyo contenido en cacao sea inferior a los 250 gramos. ¿Qué porcentaje de recipientes se espera rechazar?
- (d) ¿Mejoraría el rendimiento del proceso si la media del proceso se estableciera en 250 gramos?

6. En una fábrica de materiales de construcción, se desea estudiar si el proceso de solidificación de las vigas de hormigón se realiza bajo control estadístico. Para ello, se extrajeron muestras de cuatro vigas cada 30 minutos. Los datos de la tabla representan la presión que soporta una viga de hormigón ante determinadas condiciones ambientales:

N ^o muestra	\bar{x}	R	N ^o muestra	\bar{x}	R
1	200.00	4	13	203.50	1
2	204.00	11	14	202.20	8
3	203.00	7	15	198.00	12
4	200.75	2	16	200.25	4
5	201.00	14	17	201.50	10
6	204.00	5	18	201.00	12
7	202.00	6	19	204.50	1
8	205.00	2	20	205.00	6
9	199.75	3	21	202.75	7
10	210.50	7	22	205.00	4
11	202.00	3	23	204.25	5
12	204.00	4	24	204.25	4

- (a) ¿Qué puede deducirse respecto al funcionamiento del proceso en el análisis de los gráficos de control? Establezca los límites de vigilancia.
- (b) Si se va a controlar el proceso tomando muestras de 4 elementos cada 30 minutos, ¿cuánto tiempo se tardará en detectar un desplazamiento positivo de 3 unidades en el valor de la media del proceso?
- (c) Calcule la proporción de artículos conformes con las especificaciones si la dirección de la empresa establece un intervalo igual a 203 ∓ 1 .
- (d) Cuando ya está en marcha la producción normal, se extraen 4 nuevas muestras con 4 artículos cada una, obteniéndose los siguientes valores:

Nº muestra	\bar{x}	R
1	197	5
2	199	6
3	200	4
4	201	13

¿Qué debería hacerse con el proceso de fabricación?

7. Una fábrica de pavimentos desea controlar la duración del brillo en un nuevo material que pretende sustituir al mármol y al porcelanato. Dicho material se rocía con diversos líquidos agresivos y se mide el tiempo que el brillo tarda en desaparecer. Esta característica de calidad, distribuida normalmente, se controla por medio de dos diagramas: \bar{X} y R , y \bar{X} y S . A continuación se presentan los límites de control de dos de ellos ($n = 4$):

Diagrama de \bar{X} (basado en R)	Diagrama de S
$LSC = 626.0$	$LSC = 18.75$
$LC = 620.0$	$LC = 8.236$
$LIC = 614.0$	$LIC = 0$

Suponga que ambos diagramas indican control estadístico.

- (a) ¿Cuáles son los parámetros estimados del proceso

- i. si se utilizan los recorridos muestrales?
 - ii. si se utilizan las desviaciones muestrales?
- ¿Observa alguna anomalía?

Para responder los siguientes apartados, utilice las estimaciones obtenidas en (a.i):

- (b) Si las especificaciones para la duración del brillo son 610 ∓ 15 , ¿cuál será el porcentaje estimado de pavimento que presenta un brillo poco resistente a agentes agresivos?
 - (c) ¿Cuál es la probabilidad de detectar un cambio a 610 en la media del proceso, en la primera muestra después del cambio?
 - (d) ¿Cuál es la probabilidad de descubrir el cambio de (c) *por lo menos* en la tercera muestra después del cambio?
 - (e) Si los muestreos se realizaron cada media hora, ¿cuánto tiempo, por término medio, se necesitará para detectar el cambio de (c)?
8. En un proceso de fabricación de resistencias eléctricas cuyo valor nominal debe ser 100 ohmios, se toman muestras de 4 elementos obteniéndose los siguientes valores para las resistencias medias y las cuasidesviaciones típicas:

\bar{x}_i	s_i
101	5.1
100	6.4
110	4.3
103	5.0
96	3.4

Se pide:

- (a) ¿Es aceptable suponer que la media del proceso es realmente 100? (Indicación: utilice límites probabilísticos con $\alpha = 0.001$)
- (b) Si $\sigma = 4.5$ ohmios y se consideran defectuosas las resistencias fuera del intervalo 100 ± 8 ohmios, estime la proporción esperable de resistencias defectuosas cuando el proceso se mantiene bajo control?

- (c) Utilizando la información de (b), ¿qué produce más defectos, un aumento del 10% en la media del proceso o un aumento del 20% en la desviación típica?

Capítulo 5

Procedimientos Especiales de Control de Procesos

En este capítulo se estudian nuevos procedimientos de control de procesos:

- Comienza estudiando dos gráficos de control para características variables de calidad, que son útiles, aunque menos usuales que los estudiados con anterioridad: el gráfico de control para medidas individuales y el gráfico de control para la media y la varianza.
- Continúa con otro tipo de gráficos de control, que aprovechan la información suministrada por la sucesión de muestras anteriores.

5.1 Otros Gráficos de Control de Variables

5.1.1 Gráfico de Control para Medidas Individuales

En ocasiones, en lugar de representar la medida promedio de una muestra, \bar{x}_i , se representan cada una de las respectivas medidas individuales, es decir, el tamaño muestral utilizado es $n = 1$. Esto ocurre debido a la naturaleza de los datos o a técnicas de muestreo, como por ejemplo en la fabricación de un producto cuya terminación necesita mucho tiempo. Los gráficos en que se señalan las medidas individuales, incluso marcados los límites de especificación, son menos eficaces que los convencionales \bar{X} y R o \bar{X} y S , aunque siempre serán mejores que carecer totalmente de ellos.

En el caso de que los parámetros sean desconocidos, para la construcción de los gráficos de control se emplea la amplitud móvil de dos observaciones sucesivas para estimar la variabilidad del proceso en el gráfico R . Dados un total de m datos, las estimaciones de los parámetros del proceso son:

$$\hat{\mu} = \bar{x}; \quad \hat{\sigma} = S \text{ o bien } \hat{\sigma} = \frac{\bar{R}}{d_2}$$

donde

$$S = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2}$$

y \bar{R} se obtiene a partir de los recorridos móviles de dos observaciones consecutivas, esto es,

$$\bar{R} = \frac{1}{m-1} \sum_{i=2}^m R_i, \quad R_i = |x_i - x_{i-1}|, \quad i = 2, \dots, m.$$

Por tanto, los límites de control para el gráfico de \bar{X} dependiendo de la estimación de σ utilizada, son:

	$\hat{\sigma} = S$	$\hat{\sigma} = \frac{R}{d_2}$
LSC	$\bar{x} + kS$	$\bar{x} + \frac{k}{d_2} \bar{R}$
LC	\bar{x}	\bar{x}
LIC	$\bar{x} - kS$	$\bar{x} - \frac{k}{d_2} \bar{R}$

La constante d_2 depende del tamaño muestral que, en este caso, es 1, pero como se obtienen los R_i agrupando las observaciones de dos en dos, se puede considerar $n = 2$. Por tanto, $d_2 = 1.128$ y

$$LSC = \bar{x} + \frac{k}{1.128} \bar{R}$$

$$LC = \bar{x}$$

$$LIC = \bar{x} - \frac{k}{1.128} \bar{R}.$$

Si además $k = 3$, estos límites quedan:

$$\begin{aligned}LSC &= \bar{x} + 2.660\bar{R} \\LC &= \bar{x} \\LIC &= \bar{x} - 2.660\bar{R}.\end{aligned}$$

Los límites de control para el gráfico de R son:

$$\begin{aligned}LSC &= \bar{R} + k\frac{\bar{R}}{d_2}d_3 = \bar{R}\left(1 + k\frac{d_3}{d_2}\right) \\LC &= \bar{R} \\LIC &= \bar{R} - k\frac{\bar{R}}{d_2}d_3 = \bar{R}\left(1 - k\frac{d_3}{d_2}\right).\end{aligned}$$

Análogamente, $n = 2$, luego $d_2 = 1.128$ y $d_3 = 0.853$. Por tanto:

$$\begin{aligned}LSC &= \bar{R}\left(1 + k\frac{0.853}{1.128}\right) = \bar{R}(1 + 0.756k) \\LC &= \bar{R} \\LIC &= \bar{R}\left(1 - k\frac{0.853}{1.128}\right) = \bar{R}(1 - 0.756k).\end{aligned}$$

Si $k = 3$, entonces:

$$\begin{aligned}LSC &= D_4\bar{R} = 3.267\bar{R} \\LC &= \bar{R} \\LIC &= D_3\bar{R} = 0.\end{aligned}$$

Nota 60 Obsérvese que para medidas individuales sólo se establece un gráfico para controlar la dispersión.

Ejemplo 61 Sea un proceso de producción de un componente del motor de un coche cuya terminación requiere varias horas. Por ello, sólo se considera un componente en cada momento. Se toma una muestra de 15 componentes y se mide una característica de calidad. Las observaciones x_i obtenidas y los rangos correspondientes se muestran en la Tabla 5.1.

Las estimaciones de los distintos parámetros a partir de estos datos son:

$$\hat{\mu} = \bar{x} = 10.702, \quad \bar{R} = 1.96429, \quad \hat{\sigma} = \frac{\bar{R}}{d_2} = 1.74139$$

x_i	R_i
14.74	
12.40	2.34
8.16	4.24
12.05	3.89
10.04	2.01
9.79	0.25
10.72	0.93
9.65	1.07
12.25	2.60
9.42	2.83
8.70	0.72
11.93	3.23
10.27	1.66
9.65	0.62
10.70	1.05

Tabla 5.1: Datos individuales de la característica de calidad

y los límites de control para la media con $k = 3$, están representados en la Figura 5.1 y vienen dados por:

$$LSC = \bar{x} + 2.660\bar{R} = 15.9262$$

$$LC = \bar{x} = 10.702$$

$$LIC = \bar{x} - 2.660\bar{R} = 5.4778$$

Los límites de control para el rango con $k = 3$ están representados en la Figura 5.2 y son:

$$LSC = D_4\bar{R} = 3.267\bar{R} = 6.4205$$

$$LC = \bar{R} = 1.96429$$

$$LIC = D_3\bar{R} = 0.$$

A partir de los gráficos se puede comprobar que ningún valor muestral está fuera de los límites de control en ninguno de los dos gráficos, por lo que el proceso está bajo control en el periodo base.

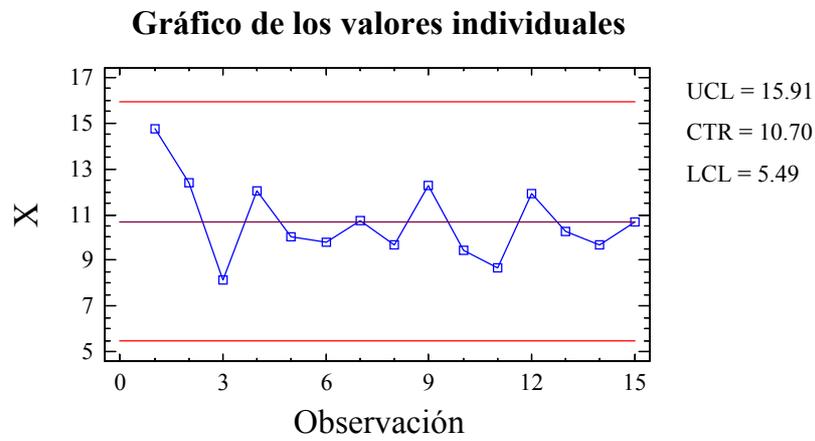


Figura 5.1: Gráfico de los valores individuales

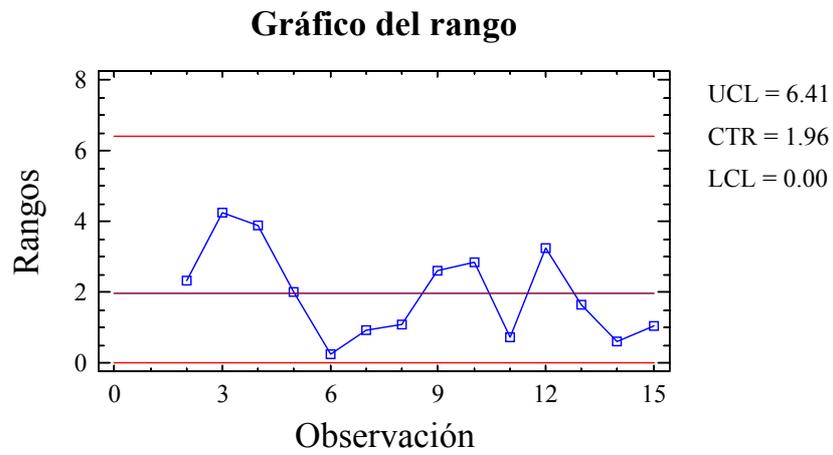


Figura 5.2: Gráfico del rango para medidas individuales

5.1.2 Gráfico de Control de Media y Varianza

La mayor parte de los estudios mediante gráficos de control para características variables de calidad se basan en los de media y rango y media y desviación típica, con preferencia del segundo para tamaños muestrales grandes. Sin embargo, algunos investigadores recomiendan la utilización de los gráficos de control de la media y varianza.

Si se extraen un total de m muestras de tamaño n cada una de ellas, se utiliza como estimador de la media poblacional la media de las medias muestrales, mientras que como estimador de la varianza se utiliza la media de la varianza de las m muestras, es decir:

$$\hat{\mu} = \bar{\bar{x}} = \frac{\sum_{i=1}^m \bar{x}_i}{m}$$

$$\hat{\sigma}^2 = \overline{S^2} = \frac{\sum_{i=1}^m S_i^2}{m}.$$

Los límites de control para el gráfico de la media si los parámetros estándares son conocidos, vienen dados por:

$$LSC = \mu + k \cdot \sigma_{\bar{X}} = \mu + k \frac{\sigma}{\sqrt{n}}$$

$$LC = \mu$$

$$LIC = \mu - k \cdot \sigma_{\bar{X}} = \mu - k \frac{\sigma}{\sqrt{n}};$$

mientras que si se tienen que estimar dichos parámetros, los límites de control se obtienen como:

$$LSC = \hat{\mu} + k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} + k \frac{\sqrt{\overline{S^2}}}{\sqrt{n}} = \bar{\bar{x}} + k \sqrt{\frac{\overline{S^2}}{n}}$$

$$LC = \hat{\mu} = \bar{\bar{x}}$$

$$LIC = \hat{\mu} - k \frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} - k \frac{\sqrt{\overline{S^2}}}{\sqrt{n}} = \bar{\bar{x}} - k \sqrt{\frac{\overline{S^2}}{n}}.$$

Los límites de control para el gráfico de la varianza no se pueden calcular de la manera habitual, $\mu_{S^2} \pm k \cdot \sigma_{S^2}$, ya que la distribución involucrada,

$\chi^2(n-1)$, no es simétrica. No obstante, se sabe que:

$$\begin{aligned} P \left[\chi_{n-1, \frac{\alpha}{2}}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{n-1, 1-\frac{\alpha}{2}}^2 \right] &= 1 - \alpha = \\ &= P \left[\frac{\sigma^2 \chi_{n-1, \frac{\alpha}{2}}^2}{(n-1)} \leq S^2 \leq \frac{\sigma^2 \chi_{n-1, 1-\frac{\alpha}{2}}^2}{(n-1)} \right] \end{aligned}$$

por lo que, utilizando la estimación de la varianza, se tiene que los límites de control vienen dados por las expresiones:

$$\begin{aligned} LSC &= \frac{\overline{S^2}}{(n-1)} \chi_{n-1, 1-\frac{\alpha}{2}}^2 \\ LC &= \overline{S^2} \\ LIC &= \frac{\overline{S^2}}{(n-1)} \chi_{n-1, \frac{\alpha}{2}}^2. \end{aligned}$$

Ejemplo 62 Sea un proceso de fabricación de un determinado producto en el que la característica de calidad que se desea estudiar es su peso. Se toman 20 muestras de tamaño 5, de las que se obtiene:

$$\begin{aligned} \hat{\mu} = \bar{\bar{x}} &= \frac{\sum_{i=1}^{20} \bar{x}_i}{20} = 49.985 \\ \hat{\sigma}^2 = \overline{S^2} &= \frac{\sum_{i=1}^{20} S_i^2}{20} = 0.56. \end{aligned}$$

Los límites de control para el gráfico de la media con $k = 3$ están representados en la Figura 5.3 y vienen dados por:

$$\begin{aligned} LSC &= \bar{\bar{x}} + k \frac{\sqrt{\overline{S^2}}}{\sqrt{n}} = 50.989 \\ LC &= \hat{\mu} = \bar{\bar{x}} = 49.985 \\ LIC &= \bar{\bar{x}} - k \frac{\sqrt{\overline{S^2}}}{\sqrt{n}} = 48.981. \end{aligned}$$

Los límites de control para la varianza con $k = 3$, es decir, $\alpha = 0.0027$,

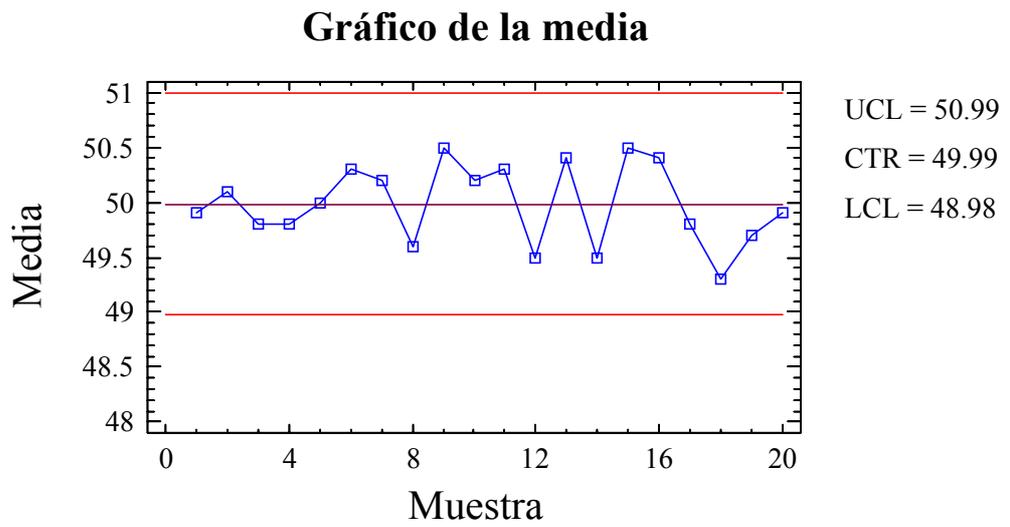


Figura 5.3: Gráfico de la media del peso del producto

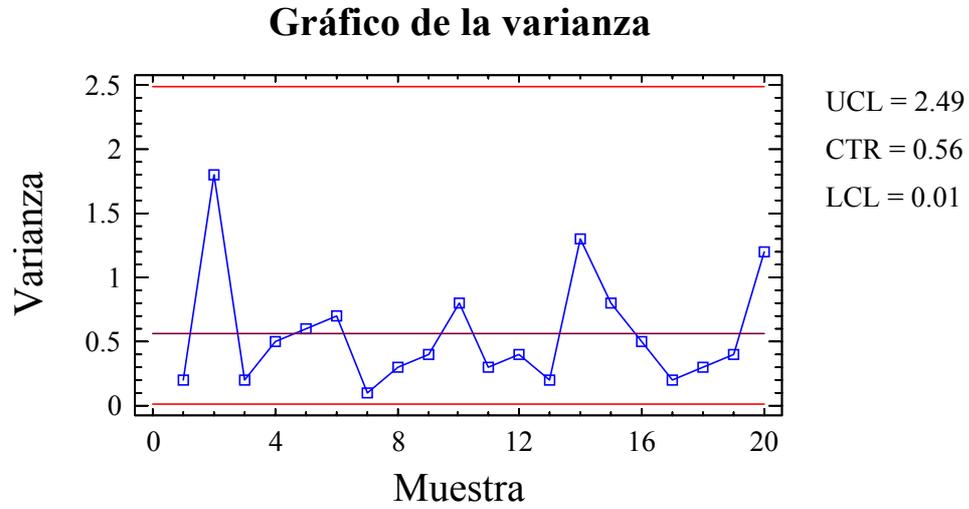


Figura 5.4: Gráfico de control de la varianza para el peso del producto

están representados en la Figura 5.4 y son:

$$LSC = \frac{\overline{S^2}}{(n-1)} \chi_{n-1, 1-\frac{\alpha}{2}}^2 = 2.4921$$

$$LC = \overline{S^2} = 0.56$$

$$LIC = \frac{\overline{S^2}}{(n-1)} \chi_{n-1, \frac{\alpha}{2}}^2 = 0.0148.$$

Se puede ver que ambos gráficos indican control estadístico en el periodo base.

5.2 Gráficos con Memoria

Hasta ahora se han estudiado los gráficos de control de Shewhart, que sólo utilizan la información contenida en el último punto y no aprovechan la información suministrada por la sucesión de puntos anteriores, por lo que son poco sensibles a pequeños cambios en la media del proceso. Por ello,

era necesario, aunque todos los puntos estuvieran bajo control estadístico, estudiar la aleatoriedad para realizar un estudio más completo. Para mejorar la capacidad para detectar cambios pequeños en la media, se introducen los gráficos de control que se indican a continuación:

- Gráfico de la media móvil (MA).
- Gráfico de la media móvil geométrica (EWMA).
- Gráficos de Sumas Acumuladas (CUSUM).

5.2.1 Gráfico de la Media Móvil (MA)

Se considera una serie de muestras independientes de tamaño n y sean $\bar{x}_1, \dots, \bar{x}_t, \dots$ las medias de las muestras respectivas. Se define la *media móvil de alcance w en el momento t* como:

$$M_t = \frac{\bar{x}_t + \bar{x}_{t-1} + \dots + \bar{x}_{t-w+1}}{w} = \frac{1}{w} \sum_{i=0}^{w-1} \bar{x}_{t-i},$$

es decir, en el instante t se suprime la media muestral más antigua y se añade la más reciente.

El gráfico de la media móvil considera que este estadístico es el que se va a representar. Para ello, se calcula su media y su varianza:

$$\mu_{M_t} = E[M_t] = E\left[\frac{1}{w} \sum_{i=0}^{w-1} \bar{x}_{t-i}\right] = \frac{1}{w} \sum_{i=0}^{w-1} E[\bar{x}_{t-i}] = \frac{1}{w} w\mu = \mu$$

$$\sigma_{M_t}^2 Var[M_t] = Var\left[\frac{1}{w} \sum_{i=0}^{w-1} \bar{x}_{t-i}\right] = \frac{1}{w^2} \left[\sum_{i=0}^{w-1} Var(\bar{x}_{t-i})\right] = \frac{1}{w^2} \left[w \frac{\sigma^2}{n}\right] = \frac{\sigma^2}{nw}.$$

Por tanto, los límites de control del gráfico *MA* vienen dados por:

$$\begin{aligned} LSC &= \mu + \frac{k\sigma}{\sqrt{nw}} \\ LC &= \mu \\ LIC &= \mu - \frac{k\sigma}{\sqrt{nw}}. \end{aligned}$$

El gráfico se completa representando M_t cuando se observe la muestra t y el alcance de la media móvil no sea mayor que el número de muestras obtenidas ($t \geq w$).

Si $1 \leq t < w$, se representa el promedio de las observaciones anteriores $\bar{x}_1, \dots, \bar{x}_t$ y los límites de control $k\sigma$ están dados por $\mu \pm \frac{k\sigma}{\sqrt{nt}}$.

Para el periodo base, se estima la media $\hat{\mu} = \bar{\bar{x}}$ y como estimación de la desviación típica se pueden utilizar las mismas que se usaron en los gráficos de control de la media y el recorrido o en la media y la desviación típica.

Nota 63 Si $w = 1$, se tiene un gráfico de Shewhart.

En este caso las medias móviles consecutivas van a estar fuertemente correlacionadas, por lo que si se tienen varios puntos consecutivos fuera de los límites de control para estabilizar el proceso en el periodo base, se debe eliminar solo el primero y recalcular los límites; así sucesivamente hasta que el proceso esté bajo control.

El gráfico *MA* es, en general, menos efectivo que el *CUSUM* y el *EWMA*, que se verán a continuación.

Ejemplo 64 Sea un proceso de fabricación de un determinado producto en el que la característica de calidad objeto de estudio es su longitud. Se toman 14 muestras de tamaño 5 y se realiza el gráfico *MA* de alcance 4, obteniendo para cada muestra los resultados que aparecen en la Tabla 5.2.

Para ilustrar el cálculo de las medias móviles se presentan aquí las cinco primeras:

$$M_1 = \bar{x}_1 = 19.68$$

$$M_2 = \frac{\bar{x}_1 + \bar{x}_2}{2} = 20.75$$

$$M_3 = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3} = 20.17$$

$$M_4 = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4}{4} = 20.27$$

$$M_5 = \frac{\bar{x}_2 + \bar{x}_3 + \bar{x}_4 + \bar{x}_5}{4} = 20.28$$

...

\bar{x}_i	R_i	M_i	z_i
19.68	0.34	19.68	20.21
21.83	0.07	20.75	20.53
19.01	0.48	20.17	20.23
20.56	0.87	20.27	20.30
19.72	1.59	20.28	20.18
19.03	0.86	19.58	19.95
22.18	0.81	20.37	20.40
21.16	0.01	20.52	20.55
20.27	0.36	20.66	20.49
21.21	0.12	21.21	20.64
19.38	0.52	20.50	20.39
19.75	1.11	20.15	20.26
19.22	0.94	19.89	20.05
21.78	0.71	20.03	20.40

Tabla 5.2: Datos sobre el peso medio, recorrido, media móvil y media móvil geométrica del producto

Los estimadores son:

$$\begin{aligned}\bar{\bar{x}} &= 20.3463 \\ \bar{R} &= 0.627919 \\ \hat{\sigma} &= \frac{\bar{R}}{d_2} = 0.269957.\end{aligned}$$

Entonces, los límites de control para $k = 3$ son:

$$\begin{aligned}LSC &= \bar{\bar{x}} + \frac{k \hat{\sigma}}{\sqrt{nt}} \\ LC &= \bar{\bar{x}} = 20.3463 \\ LIC &= \bar{\bar{x}} - \frac{k \hat{\sigma}}{\sqrt{nt}}\end{aligned}$$

para las $t < w$ primeras muestras, es decir,

$$\begin{array}{lll}t = 1 & LSC = 20.7084, & LIC = 19.9841 \\ t = 2 & LSC = 20.6024, & LIC = 20.0902 \\ t = 3 & LSC = 20.5554, & LIC = 20.1372\end{array}$$

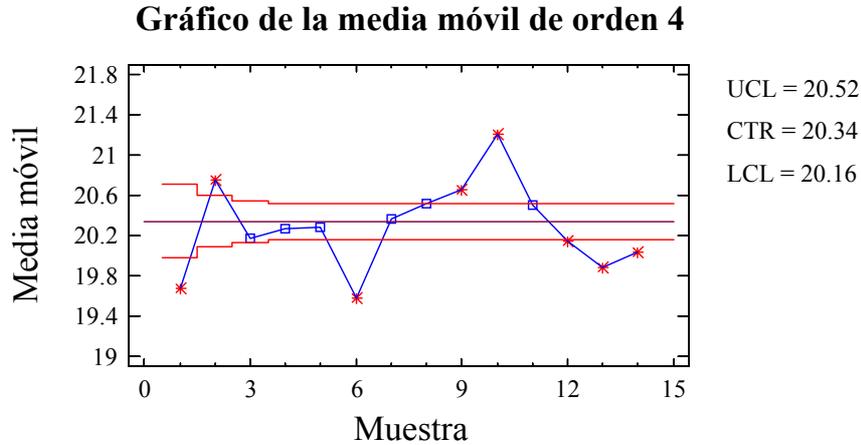


Figura 5.5: Gráfico de la media móvil de orden 4 para la longitud del producto

y a partir de la muestra 4,

$$LSC = \bar{\bar{x}} + \frac{k \hat{\sigma}}{\sqrt{nw}} = 20.5274$$

$$LC = \bar{\bar{x}} = 20.3463$$

$$LIC = \bar{\bar{x}} - \frac{k \hat{\sigma}}{\sqrt{nw}} = 20.1652.$$

En la Figura 5.5 está representado el gráfico de control donde se puede observar 7 muestras fuera de los límites de control. Se concluye que el proceso está fuera de control ya que su media es inestable en el periodo base.

5.2.2 Gráfico de la Media Móvil Geométrica (EWMA)

En el gráfico de la media móvil, las w medias tenían la misma ponderación, $\frac{1}{w}$, y el resto tenían peso nulo, es decir, se da el mismo peso a cada media sin considerar cuanto tiempo pasó desde que se obtuvo la información hasta el instante considerado. En el gráfico *EWMA* se considera

una ponderación en la que los pesos decrecen exponencialmente con el paso del tiempo, es decir, se asigna un peso $r \leq 1$ a la muestra actual t , pesos $r(1-r)$, $r(1-r)^2, \dots, r(1-r)^t$ a las muestras $(t-1)$, $(t-2), \dots, 1$, respectivamente y peso $(1-r)^t$ al valor de partida $\bar{x}_0 = \mu$, de manera que la suma de todos los pesos es 1^1 .

Por tanto, el estadístico objeto del gráfico de control es:

$$\begin{aligned} z_t &= r \bar{x}_t + (1-r)z_{t-1} = \\ &= r \bar{x}_t + r(1-r)\bar{x}_{t-1} + r(1-r)^2\bar{x}_{t-2} + \dots + r(1-r)^{t-1}\bar{x}_1 + (1-r)^t\bar{x}_0, \end{aligned}$$

donde $z_0 = \bar{\bar{x}}$, $0 < r \leq 1$, $t = 1, 2, \dots$; esto es, en cada momento t , se pondera la observación actual \bar{x}_t con el valor de la media móvil anterior z_{t-1} .

Para su representación, se calcula su media y su varianza supuesto que hay estabilidad en la media e independencia entre las muestras:

$$\begin{aligned} \mu_{z_t} &= E[z_t] = r\mu + r(1-r)\mu + r(1-r)^2\mu + \dots + r(1-r)^{1-r}\mu + (1-r)^t\mu = \\ &= \mu (r + r(1-r) + r(1-r)^2 + \dots + r(1-r)^{1-r} + (1-r)^t) = \mu \end{aligned}$$

$$\begin{aligned} Var[z_t] &= Var[r \bar{x}_t + r(1-r)\bar{x}_{t-1} + \dots + r(1-r)^{t-1}\bar{x}_1 + (1-r)^t\bar{x}_0] = \\ &= r^2\frac{\sigma^2}{n} + r^2(1-r)^2\frac{\sigma^2}{n} + r^2(1-r)^4\frac{\sigma^2}{n} + \dots + r^2(1-r)^{2(t-1)}\frac{\sigma^2}{n} = \\ &= \frac{\sigma^2}{n} (r^2 + r^2(1-r)^2 + r^2(1-r)^4 + \dots + r^2(1-r)^{2(t-1)}) = \\ &= \frac{\sigma^2}{n} r^2 \left(\frac{1 - (1-r)^{2t}}{1 - (1-r)^2} \right) = \frac{\sigma^2}{n} \left(\frac{r}{2-r} \right) (1 - (1-r)^{2t}). \end{aligned}$$

Por tanto, los límites de control del gráfico *EWMA* son:

$$\begin{aligned} LSC &= \mu + \frac{k\sigma}{\sqrt{n}} \sqrt{\frac{r}{2-r}} \sqrt{1 - (1-r)^{2t}} \\ LC &= \mu \\ LIC &= \mu - \frac{k\sigma}{\sqrt{n}} \sqrt{\frac{r}{2-r}} \sqrt{1 - (1-r)^{2t}} \end{aligned}$$

¹ $r \sum_{k=0}^{t-1} (1-r)^k + (1-r)^t = r \frac{(1-r)^t - 1}{1-r-1} + (1-r)^t = 1 - (1-r)^t + (1-r)^t = 1$

que tienden a estabilizarse, cuando t aumenta², a

$$\begin{aligned}LSC &= \mu + \frac{k\sigma}{\sqrt{n}} \sqrt{\frac{r}{2-r}} \\LC &= \mu \\LIC &= \mu - \frac{k\sigma}{\sqrt{n}} \sqrt{\frac{r}{2-r}}.\end{aligned}$$

Nota 65 Si se considera $r = 1$, se tiene un gráfico Shewhart, y si $r = \frac{2}{w+1}$, se tiene un gráfico MA.

Por lo general, se suelen tomar valores de r situados entre 0.05 y 0.25 ($r = 0.05, 0.10, 0.20, 0.25$), de manera que cuanto mayor sea su valor, menor importancia se le da a los valores más alejados en el tiempo.

Para obtener los límites de control en el periodo base se estima la media mediante $\hat{\mu} = \bar{\bar{x}}$ y la desviación típica como en el gráfico MA.

Este gráfico es más eficaz que el MA y el de la media para detectar pequeños cambios en el proceso.

Ejemplo 66 Con los datos del Ejemplo 64 se quiere construir un gráfico de medias móviles geométricas con $r = 0.2$.

Las medias móviles geométricas para $r = 0.2$ se calculan como se indica a continuación:

$$\begin{aligned}z_0 &= \bar{\bar{x}} = 20.3463 \\z_1 &= 0.2\bar{x}_1 + 0.8z_0 = 20.21 \\z_2 &= 0.2\bar{x}_2 + 0.8z_1 = 20.53 \\&\dots\end{aligned}$$

y aparecen en la Tabla 5.2.

Entonces, los límites de control obtenidos cuando t aumenta son:

$$\begin{aligned}LSC &= 20.467 \\LC &= 20.3463 \\LIC &= 20.2256.\end{aligned}$$

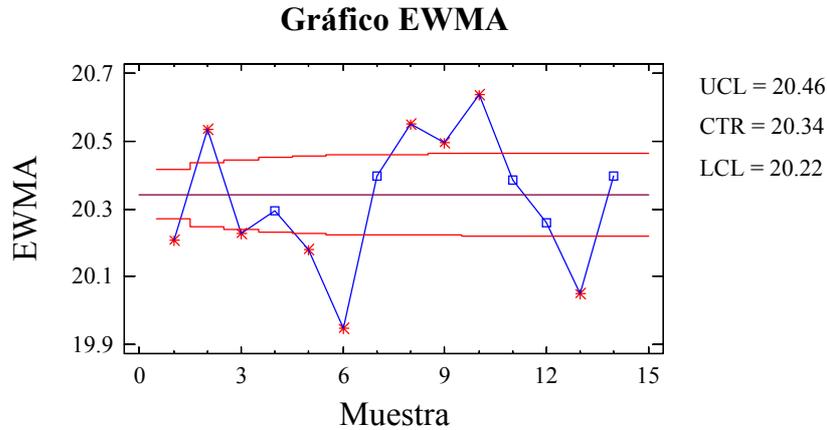


Figura 5.6: Gráfico *EWMA* con $r = 0.2$ para la longitud

En la Figura 5.6 se representa el gráfico *EWMA* y se puede comprobar que 9 muestras se salen de los límites de control, por lo que se concluye que la media del proceso es inestable y está fuera de control en el periodo base.

5.2.3 Gráficos de Sumas Acumuladas (*CUSUM*)

Uno de los métodos posibles para el tratamiento de la información incluida en las anteriores observaciones, es utilizar sólo la última secuencia de observaciones que pueda indicar un cambio de μ . Los gráficos *CUSUM* se basan en la representación de la acumulación de las desviaciones de cada observación respecto a un valor de referencia. El término *CUSUM* procede del inglés, *Cumulative Sum* (suma acumulada).

El gráfico *CUSUM* es más sensible que el gráfico de Shewhart para detectar pequeños cambios en el proceso, pero es más lento para detectar cambios grandes.

En las siguientes secciones se muestran dos métodos de construcción de un gráfico *CUSUM*:

$${}^2t \rightarrow \infty, r < 1 \implies (1-r)^{2t} \rightarrow 0$$

Construcción de un Gráfico CUSUM: Método I

Se pueden detectar dos tipos de cambios en la media del proceso:

- Si se quiere detectar sólo un cambio a valores mayores de la media del proceso, el contraste es:

$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu = \mu_1 \end{array} \right\}$$

con $\mu_0 < \mu_1$.

Para resolver el contraste, se extraen m muestras y sea \bar{X}_i la media del subgrupo i -ésimo. Cuando el proceso está bajo control y H_0 es cierta, la variable aleatoria $Z_i = \bar{X}_i - \mu_0$ tiene media igual a 0. Además, la variable aleatoria $Y_i = Z_i - d$, con $d > 0$, tendrá esperanza negativa e igual a $-d$. La suma acumulada de Y_i ,

$$S_k = \sum_{i=1}^k Y_i,$$

tenderá a decrecer mientras que H_0 sea cierta, ya que $E[S_1] = -d$, $E[S_2] = -2d$, ..., $E[S_k] = -kd$.

Si el proceso pasa a una situación *fuera de control* y H_1 es cierta, la esperanza de Y_i ya no será $-d$, sino:

$$E[Y_i|H_1] = E[\bar{X}_i - \mu_0 - d|H_1] = E[\bar{X}_i|H_1] - \mu_0 - d = \mu_1 - \mu_0 - d$$

Como $\mu_0 < \mu_1 \Rightarrow \mu_1 - \mu_0 > 0$, se puede tomar un valor de d tal que $E[Y_i|H_1] > 0$. Se recomienda tomar el valor:

$$d = \frac{\mu_1 - \mu_0}{2}.$$

Entonces, la suma acumulada tenderá a aumentar, mientras que H_1 sea cierta.

Un gráfico *CUSUM* es la representación de la suma acumulada S_k . En tanto que S_k esté generalmente disminuyendo, se deja sólo el proceso y se continúa el muestreo. Cuando se produce un cambio *significativo* hacia arriba en la marca de S_k , se da una señal de que el proceso está *fuera de control*; es decir, si el valor más reciente de S_k excede en c

unidades a un nivel previo de la suma acumulada, se da una señal, donde la elección de la constante c es un tanto arbitraria (se suele tomar $c = 6d$). En la práctica, se compara la suma acumulada S_k con

$$c + \min_{i=1, \dots, k} S_i$$

y, mientras la suma acumulada se mantenga por debajo de esta cantidad, el proceso está bajo control. En caso contrario, el proceso pasa a un estado de fuera de control.

- En el caso en que sea necesario detectar sólo un cambio a valores menores de la media del proceso, esto es

$$\left. \begin{array}{l} H_0 : \mu = \mu_0 \\ H_1 : \mu = \mu_1 \end{array} \right\}$$

con $\mu_0 > \mu_1$, hay que tener en cuenta lo siguiente:

Ahora, bajo H_0 , la variable aleatoria $Y_i = Z_i - d$, con $d < 0$, tendrá esperanza positiva e igual a $-d$. La suma acumulada de Y_i ,

$$S_k = \sum_{i=1}^k Y_i,$$

tenderá a moverse hacia arriba mientras que H_0 sea cierta.

Si el proceso pasa a una situación *fuera de control* y H_1 es cierta, la esperanza de Y_i ya no será $-d$, sino, al igual que en el caso anterior, $\mu_1 - \mu_0 - d$. Por otra parte, como $\mu_0 > \mu_1 \Rightarrow \mu_1 - \mu_0 < 0$, y

$$d = \frac{\mu_1 - \mu_0}{2} < 0,$$

entonces, la suma acumulada de Y_i tenderá a moverse hacia abajo, mientras que H_1 sea cierta.

De este modo, mientras que S_k esté generalmente aumentando, se deja sólo el proceso y se continúa el muestreo. Cuando se produce un cambio *significativo* hacia abajo en la marca de S_k , se da una señal de que el proceso está *fuera de control*; es decir, si el valor más reciente de S_k es inferior en c unidades (usualmente $6d$) a un nivel previo de

la suma acumulada se da una señal. En la práctica, se compara la suma acumulada S_k con:

$$c + \max_{i=1, \dots, k} S_i$$

y, mientras la suma acumulada se mantenga por encima de esta cantidad, el proceso está bajo control.

Construcción de un Gráfico CUSUM: Método II

Este método de construcción del gráfico *CUSUM* se basa en el uso de la *maskarilla V* propuesta por Barnhard (1959).

Se quiere detectar un cambio en la media del proceso de una magnitud Δ , esto es:

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &= \mu_0 \pm \Delta \end{aligned}$$

Se extraen m muestras y sea \bar{X}_i la media del subgrupo i -ésimo. El estadístico que se representa es la suma acumulada de orden k , que viene dado por la expresión:

$$S_k = \sum_{i=1}^k (\bar{X}_i - \mu_0)$$

e incorpora toda la información de la sucesión de valores muestrales anteriores. Bajo la hipótesis nula, $E[S_m] = 0$.

El procedimiento de decisión consiste en aplicar la *maskarilla V* al diagrama de suma acumulada, con el punto O en el último valor, S_m , y la línea OA paralela al eje horizontal de longitud d (véase Figura 5.7). Si todas las sumas acumuladas S_1, \dots, S_m caen entre los dos brazos de la *maskarilla*, el proceso se encuentra bajo control, puesto que S_k varía aleatoriamente alrededor de la línea cero. Sin embargo, si cualquier S_k se sitúa fuera de los brazos de la *maskarilla*, el proceso se considera fuera de control. Específicamente, si el punto se sitúa por encima del brazo superior, ello indicará un cambio en la media a valores menores ($\mu_0 - \Delta$); por el contrario, si el punto se sitúa por debajo del brazo inferior, reflejará un cambio en la media a valores mayores ($\mu_0 + \Delta$).

Los parámetros de la mascarilla V indican el funcionamiento del gráfico *CUSUM*. Se utiliza la *distancia guía* d y el ángulo θ para definir dicha mascarilla:

$$d = \frac{2}{\delta^2} \ln \left(\frac{1-\beta}{\alpha} \right), \quad \theta = \arctan \left(\frac{\Delta}{2k} \right), \quad \delta = \frac{\Delta}{\sigma_{\bar{X}}}$$

donde

$$\alpha = P[\text{Error I}]$$

$$\beta = P[\text{Error II}]$$

Δ = cambio en la media del proceso que se quiere detectar

δ = magnitud del cambio que se quiere detectar

y k es un factor de escala, que relaciona la unidad de escala vertical con la unidad de escala horizontal, habitualmente $\sigma_{\bar{X}} \leq k \leq 2\sigma_{\bar{X}}$

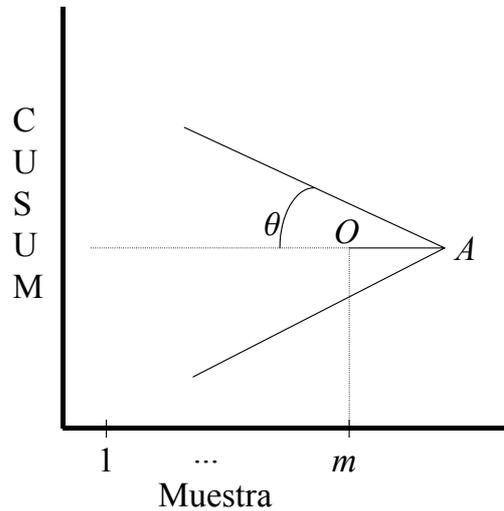


Figura 5.7: Gráfico CUSUM con Mascarilla V

Nota 67 *El gráfico EWMA es menos sensible que el CUSUM a pequeñas desviaciones. Por esto, se recomienda para mejorar el comportamiento de los gráficos EWMA ante pequeñas desviaciones (inferiores a 1σ) utilizar k entre 2.6 y 2.8 si $r \leq 0.1$.*

$N^{\circ}muestra$	\bar{x}_i	$N^{\circ}muestra$	\bar{x}_i
1	20	16	37
2	56	17	97
3	39	18	17
4	42	19	80
5	77	20	95
6	21	21	96
7	27	22	96
8	23	23	102
9	12	24	95
10	100	25	90
11	97	26	76
12	37	27	88
13	98	28	90
14	60	29	65
15	12	30	102

Tabla 5.3: Datos de la longitud media por muestra de tornillos

Ejemplo 68 *La medida de un lápiz es de 60 mm. El largo historial de fabricación de este tipo de lápices muestra que la desviación típica es de 28 mm. El proceso se consideraría insatisfactorio si la longitud media creciera hasta 80 mm. Se pretende poder detectar este cambio desde 60 mm. hasta 80 mm. Se toman 30 muestras de tamaño 4 de este producto, de las que se obtienen sus medias, que se encuentran en la Tabla 5.3.*

- *Método I:*

Se toma $d = \frac{\mu_1 - \mu_0}{2} = \frac{80 - 60}{2} = 10$ y $c = 6 \cdot d = 60$.

Se tabulan los valores de $Y_i = \bar{X}_i - \mu_0 - d$ y de $S_k = \sum_{i=1}^k Y_i$ en la Tabla 5.4. En la Figura 5.8 se observa el gráfico CUSUM, que señala el cambio en la muestra 21. A partir de este punto el gráfico se mantiene por encima del nivel de la señal. Entonces, se puede decir que se ha producido un cambio en la media a partir de la muestra 21, de manera que a partir de esa muestra, la media es mayor o igual que 80.

- *Método II:*

<i>Muestra</i>	\bar{x}_i	Y_i	S_i	$c + \text{mín}S_i$	<i>Muestra</i>	\bar{x}_i	Y_i	S_i	$c + \text{mín}S_i$
1	20	-50	-50	+10	16	37	-33	-368	-308
2	56	-14	-64	-4	17	97	+27	-341	-308
3	39	-31	-95	-35	18	17	-53	-394	-334
4	42	-28	-123	-63	19	80	+10	-384	-334
5	77	+7	-116	-63	20	95	+25	-359	-334
6	21	-49	-165	-105	21	96	+26	-333	-334
7	21	-49	-214	-154	22	96	+26	-307	-334
8	23	-47	-261	-201	23	102	+32	-275	-334
9	12	-58	-319	-259	24	95	+25	-250	-334
10	100	+30	-289	-259	25	90	+20	-230	-334
11	97	+27	-262	-259	26	76	+6	-224	-334
12	37	-33	-295	-259	27	88	+18	-206	-334
13	98	+28	-267	-259	28	90	+20	-186	-334
14	60	-10	-277	-259	29	65	-5	-191	-334
15	12	-58	-335	-275	30	102	+32	-159	-334

Tabla 5.4: Resolución del gráfico CUSUM mediante el método I

Se tabulan los valores de $Y_i = \bar{X}_i - \mu_0$ y de $S_k = \sum_{i=1}^k Y_i$ en la Tabla 5.5.

Si se considera $\alpha = 0.0027$, $\beta = 0.01$, $\Delta = 20$, $k = 2\sigma_{\bar{X}} = 2\frac{28}{\sqrt{4}} = 28$,

$$\delta = \frac{\Delta}{\sigma_{\bar{X}}} = \frac{20}{14} = 1.428, \theta = \arctan\left(\frac{\Delta}{2k}\right) = 19.65^\circ,$$

$$d = \frac{2}{\delta^2} \ln\left(\frac{1-\beta}{\alpha}\right) = 5.7863$$

Entonces, en la Figura 5.9 se representa el gráfico CUSUM mediante el Método II, que indica que se ha producido un cambio en la media de más de 20 mm. hacia arriba.

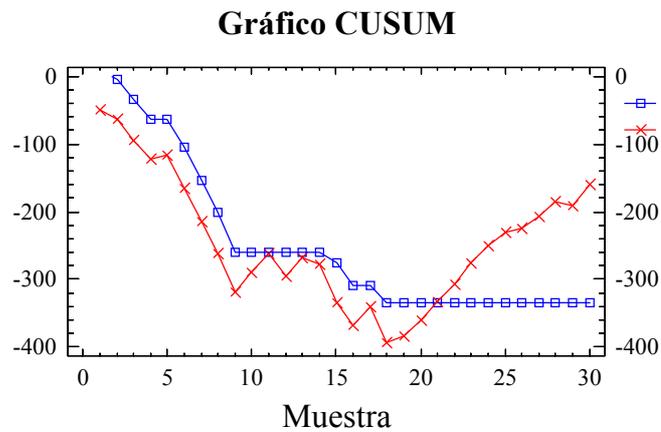


Figura 5.8: Gráfico *CUSUM* mediante el Método I

5.3 Ejercicios

1. Se realizaron las siguientes lecturas sobre el consumo diario de botellas de vino tinto en un restaurante:

Día	Consumo	Día	Consumo
1	95.5	16	93
2	96.5	17	87.5
3	90.5	18	96.5
4	100	19	98
5	97	20	96
6	93.5	21	98.5
7	90	22	94
8	89.5	23	96.5
9	92.5	24	98
10	96.5	25	92
11	97	26	93.8
12	99.5	27	96
13	91	28	94
14	102	29	95
15	87	30	95.4

Realice un gráfico de control para medidas individuales.

<i>Muestra</i>	\bar{x}_i	Y_i	S_i	<i>Muestra</i>	\bar{x}_i	Y_i	S_i
1	20	-40	-40	16	37	-23	-208
2	56	-4	-44	17	97	+37	-171
3	39	-21	-65	18	17	-43	-214
4	42	-18	-83	19	80	+20	-194
5	77	+17	-66	20	95	+35	-159
6	21	-39	-105	21	96	+36	-123
7	21	-39	-144	22	96	+36	-87
8	23	-37	-181	23	102	+62	-45
9	12	-48	-229	24	95	+35	-10
10	100	+40	-189	25	90	+30	+20
11	97	+37	-152	26	76	+16	+36
12	37	-23	-175	27	88	+28	+64
13	98	+38	-137	28	90	+30	+94
14	60	0	-137	29	65	+5	+99
15	12	-48	-185	30	102	+42	+141

Tabla 5.5: Resolución del gráfico CUSUM mediante el Método II

2. Se estudia el contenido de los paquetes de azúcar de una marca determinada. Se analizan 15 muestras de tamaño 10 y, tras el correspondiente cambio de variable para facilitar los cálculos, se obtienen

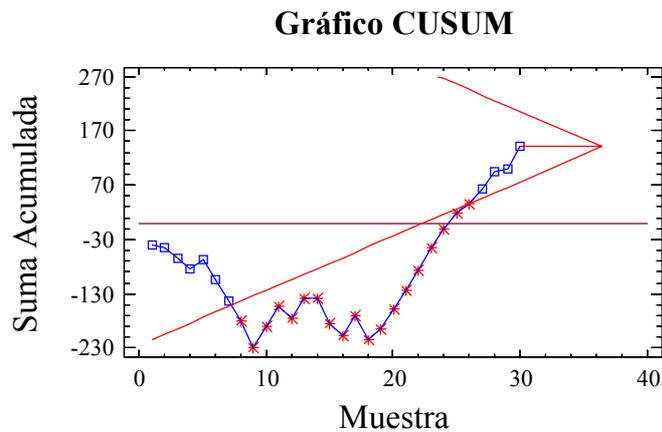


Figura 5.9: Gráfico CUSUM mediante el Método II

los siguientes resultados:

Muestra\Dato	1	2	3	4	5	6	7	8	9	10
1	2.5	0.5	2	-1	1	-1	0.5	1.5	0.5	-1.5
2	0	0	0.5	1	1.5	1	-1	1	1.5	-1
3	1.5	1	1	-1	0	-1.5	-1	-1	1	-1
4	0	0.5	-2	0	-1	1.5	-1.2	0	-2	-1.5
5	0	0	0	-0.5	0.5	1	-0.5	-0.5	0	0
6	1	-0.5	0	0	0	0.5	-1	1	-2	1
7	1	-1	-1	1	0	1.5	0	1	0	0
8	0	-1.5	-0.5	1.5	0	0	0	-1	0.5	-0.5
9	-2	-1.5	1.5	1.5	0	0	0.5	1	0	1
10	-0.5	3.5	0	-1	-1.5	-1.5	-1	-1	1	0.5
11	0	1.5	0	0	2	-1.5	0.5	-0.5	2	-1
12	0	-2	-0.5	0	-0.5	2	1.5	0	0.5	-1
13	-1	-0.5	-0.5	-1	0	0.5	0.5	-1.5	-1	-1
14	0.5	1	-1	-0.5	-2	-1	-1.5	0	1.5	1.5
15	1	0	1.5	1.5	1	-1	0	1	-2	-1.5

Elabore los gráficos de control de media y varianza para el periodo base del proceso.

- Se estudia la cantidad de refresco de cola que hay en una botella de 500 ml. Las latas se llenan en una máquina automática. Cada 30

minutos de fabricación se toman muestras de 5 latas con los siguientes resultados:

Hora	Media	Recorrido	Hora	Media	Recorrido
1	495.5	12.8	13	489	15.2
2	493	13.2	14	493.5	11.6
3	496	15.3	15	488.4	11.8
4	497.5	8.3	16	495.5	7.6
5	490.5	7.4	17	491	8.4
6	496	8.8	18	497	7.3
7	500	13.2	19	494.5	8.7
8	498	7.3	20	490.5	9.8
9	497	8.4	21	487.5	9.8
10	496.6	9.2	22	489.4	11.2
11	493.4	10.2	23	488	10.8
12	498	12.3	24	486.5	12.5

- (a) Represente el gráfico de la media móvil de orden 4.
 (b) Represente el gráfico de la media móvil geométrica con $r = 0.5$.
4. Los siguientes datos representan la dureza media y el recorrido de 30 muestras de 5 cascos de plástico tomados en cada hora de producción:

Muestra	Media	Recorrido	Muestra	Media	Recorrido
1	29	11	16	23.8	4
2	25	8	17	27	5
3	26	5	18	28.4	6
4	25.2	5	19	25.4	7
5	25.4	3	20	26.2	6
6	28	4	21	27	5
7	26	5	22	26	5
8	27	4	23	28	3
9	24.8	7	24	26.4	6
10	21.4	4	25	27.3	4
11	23.9	3	26	24	4
12	24.1	5	27	22.6	2
13	27	4	28	28	6
14	26.8	6	29	24.4	5
15	26.4	2	30	26.4	3

- (a) Obtenga los límites de control para la media y el recorrido.
- (b) Se considera negativo que el promedio de la dureza supere el valor 27.5. Construya un gráfico CUSUM con $c = 6d$ para realizar el contraste:

$$\left. \begin{array}{l} H_0 : \mu = 25.90 \\ H_1 : \mu = 27.50 \end{array} \right\}$$

- (c) Repita b) con la hipótesis alternativa $H_1 : \mu = 24.30$.

Capítulo 6

Análisis de la Capacidad del Proceso

6.1 Introducción

Como ya se ha comentado anteriormente, aún cuando el proceso está bajo control estadístico, es decir, está trabajando bajo un sistema de causas fortuitas, es imposible que dos unidades de fabricación sean idénticas, siempre hay una variabilidad entre las unidades de fabricación.

Una característica de calidad siempre tiene un valor óptimo que se conoce con el nombre de *valor nominal* u *objetivo*. Dado que existe una variabilidad natural, cuando se diseña un producto no se determina un único valor objetivo, sino que también se especifica un intervalo, llamado *intervalo de tolerancia* o *de especificación*. Si la característica de calidad observada en una unidad de fabricación está dentro de dicho intervalo se considera que la unidad es aceptable, de otro modo se rechaza.

En este capítulo estamos interesados en medir la *capacidad del proceso* de cumplir con las especificaciones, es decir, estamos interesados en medir la aptitud de un proceso para producir unidades dentro de los límites de especificación.

Veamos los elementos que van a ser necesarios para juzgar la capacidad del proceso:

- En primer lugar el proceso ha de estar bajo control estadístico, en otro caso no tiene sentido evaluar la capacidad del proceso para cumplir

con los requerimientos impuestos por la administración.

- Se han de conocer los *límites de especificación*, que son los extremos del intervalo de tolerancia. Estos límites deben ser fijados teniendo en cuenta aspectos funcionales, comerciales, etc. y no la variabilidad natural del proceso.
- Datos del proceso que describan su funcionamiento. Nótese que, puesto que se parte de un proceso bajo control estadístico, ya se dispone de datos: las muestras utilizadas en los diagramas de control establecidos para vigilar el proceso.
- Estimaciones de la media y variabilidad del proceso. Una vez que se dispone de datos, estos valores han de ser trasladados a estimaciones sobre el comportamiento del proceso, entre ellas la media y la variabilidad del mismo.

Una de las suposiciones más habituales es la de que los datos provienen de una distribución normal, es por ello que la mayoría de las medidas de capacidad asumen tácitamente la distribución normal en sus definiciones. Esto es algo a tener en cuenta, de modo que cuando tratemos de establecer la capacidad de un proceso, hemos de considerar que algunos no siguen una distribución normal y debemos estudiar los datos para determinar qué distribución es mejor utilizar.

6.2 Estimación de la Variabilidad del Proceso

Dado que uno de los requerimientos para estudiar la capacidad del proceso es que este se encuentre bajo control estadístico, podemos utilizar la información que nos proporcionan los diagramas de control establecidos para conseguir dicho control. Así, para vigilar la variabilidad de características de calidad que sean variables, se utilizan los diagramas de R o S y, por lo tanto, disponemos de la estimación de la desviación típica del proceso como

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \text{ ó } \hat{\sigma} = \frac{\bar{S}}{c_4}$$

Estas dos estimaciones suponen que la distribución subyacente es normal. En los casos en los que dicha distribución no sea normal se necesitarán otras estimaciones de σ .

Otro método para estimar la variación del proceso es ignorar el hecho de que los datos están agrupados en muestras en el diagrama de control, y estimar la desviación típica del proceso utilizando la desviación típica muestral de todos los datos. Dicha estimación es tan buena como las anteriores en el caso de que el proceso esté bajo control pero, en el caso de que no sea así, será mayor, ya que reflejará no sólo la variación dentro de los subgrupos sino también la variación entre subgrupos.

Para estimar la media del proceso utilizaremos la línea central del diagrama de \bar{X} .

Una vez estimadas las media y la desviación típica del proceso $(\hat{\mu}, \hat{\sigma})$, se puede obtener una medida de la *variabilidad natural* del proceso, que se cuantifica como la longitud del intervalo $\hat{\mu} \pm 3\hat{\sigma}$, esto es, $6\hat{\sigma}$. Esta definición proviene de suponer que la distribución subyacente es Normal. En el caso de no ser así, se puede definir la variabilidad natural del proceso como la amplitud del intervalo centrado en la media poblacional y que contiene el 99.73% de los valores del proceso.

6.3 Límites de Control, Límites de Tolerancia Natural y Límites de Especificación

Una vez que se ha introducido el concepto de límite de especificación, creemos que es un buen momento para aclarar las diferencias entre los distintos tipos de límites que han sido introducidos hasta ahora, en particular los límites de tolerancia natural, los límites de control y los límites de especificación:

- Límites de tolerancia natural, son los extremos del intervalo en el que toma valores la variable del proceso, X . Estos son motivados por la variabilidad natural del proceso, en particular, si $X \rightsquigarrow N(\mu, \sigma)$, los límites de tolerancia natural son $\mu \pm 3\sigma$.
- Límites de control, son los extremos del intervalo en el que toma valores el estadístico, W , seleccionado para controlar el proceso, en otras palabras, son los límites de tolerancia natural del estadístico W . Por ejemplo, si para controlar la media del proceso se va a utilizar un diagrama de \bar{X} , los límites de control son $\mu \pm 3\sigma/\sqrt{n}$.

- Límites de especificación, son límites que son determinados externamente, sin tener en cuenta la variabilidad del proceso, y que pueden estar impuestos por la administración, el cliente o los diseñadores del producto.

Nótese que los límites de especificación están relacionados con la variable del proceso X , y no con la \bar{X} , por lo tanto no se deben dibujar en el diagrama de control de \bar{X} . Estos se pueden representar en el diagrama de control sólo cuando se trabaja con observaciones individuales y no con medias.

Ejemplo 69 *Considérese el ejemplo la compañía de seguros del Capítulo 4. El diagrama de control de la \bar{X} se estableció con los siguientes límites de control:*

$$LSC = 110.46$$

$$LIC = 47.71$$

La línea central del diagrama de \bar{X} fue $LC = 77.58$, que coincide con la estimación de la media del proceso. Por otra parte, la estimación de la desviación típica del proceso fue:

$$\hat{\sigma} = 24.51$$

con lo que si suponemos normalidad en los datos, $X \rightsquigarrow N(77.58, 24.51)$. Los límites de tolerancia natural son por tanto:

$$LSTN = \hat{\mu} + 3\hat{\sigma} = 77.58 + 3 \cdot 24.51 = 151.11$$

$$LITN = \hat{\mu} - 3\hat{\sigma} = 77.58 - 3 \cdot 24.51 = 4.05$$

que como se puede observar son más amplios que los límites de control.

La administración de la compañía de seguros decide que no se hará cargo de las primas que superen los 3000 euros. Esta decisión establece un límite de especificación superior de 300, que como se puede observar no tiene que ver ni con la variabilidad natural del proceso, ni con la variabilidad del estadístico seleccionado para su control (\bar{X}).

6.4 Estimación de la Proporción de No Conformes

Después de estimar los parámetros del proceso y conocer así su funcionamiento, podemos estimar la proporción de elementos que no cumplen con las especificaciones, ya sea porque caen por encima del límite superior de especificación (LSE) o porque caen por debajo del límite inferior de especificación (LIE). En otras palabras, vamos a estimar la proporción de elementos no conformes.

Si suponemos que la distribución subyacente del proceso es $N(\mu, \sigma)$ entonces las proporciones de no conformes se estiman por:

$$\begin{aligned} \text{Proporción sobre } LSE &= P(X > LSE) = P\left(Z > \frac{LSE - \hat{\mu}}{\hat{\sigma}}\right) \\ \text{Proporción bajo } LIE &= P(X < LIE) = P\left(Z < \frac{LIE - \hat{\mu}}{\hat{\sigma}}\right), \end{aligned}$$

donde $Z \sim N(0, 1)$.

Si la distribución no es Normal, hay que calcular $P(X > LSE)$ y $P(X > LIE)$ utilizando la distribución de probabilidad de la variable X .

Ejemplo 70 *Un proceso de fabricación bajo control estadístico produce bombillas con una duración media estimada de 150 h. Se han estudiado datos históricos del proceso y se sabe que su distribución es exponencial. Si el límite inferior de especificación es de 70 horas, la proporción de bombillas no conformes que se están produciendo se calcula como:*

$$\begin{aligned} X \sim \exp(\lambda) &\Rightarrow E(X) = \frac{1}{\lambda} = 150 \Rightarrow \lambda = \frac{1}{150} \\ P(X < LIE) &= F(LIE) = 1 - e^{-\frac{LIE}{150}} = 1 - e^{-\frac{70}{150}} = 0.37291. \end{aligned}$$

Si hubiésemos considerado la distribución normal:

$$P(X < LIE) = P\left(Z < \frac{LIE - 150}{150}\right) = P(Z < -0.533) = 0.2969.$$

6.5 Índices de Capacidad

En la sección anterior hemos visto como evaluar la capacidad del proceso calculando la proporción de elementos que se salen de los límites de

especificación. Otra forma de evaluar dicha capacidad es calcular medidas o *índices de capacidad* del proceso. En dichos índices aparece la amplitud del proceso que hemos definido anteriormente, la cual está inspirada en la distribución normal, por lo que antes de calcular e interpretar dichos índices, habrá que comprobar si efectivamente estamos trabajando con una distribución normal.

1. Considera los dos procesos siguientes (el tamaño muestral $n = 5$)

Proceso A	Proceso B
$\bar{x}_A = 100$	$\bar{x}_B = 105$
$\bar{S}_A = 3$	$\bar{S}_B = 1$

Las especificaciones son 100 ± 10 . Calcular todos los índices de capacidad e interpretarlos. ¿Qué proceso prefieres? Los Índices C_p , C_{pu} , C_{pl} y C_{pk}

Índice de Capacidad del Proceso C_p

Con este índice comparamos la variabilidad natural del proceso con la longitud del intervalo de especificación. Comentamos anteriormente que la amplitud natural del proceso se cuantifica por 6σ . Por otra parte, la longitud del intervalo entre los límites de especificación sería el rango de valores que permitimos al proceso. El *índice de capacidad del proceso*, C_p , se define por

$$C_p = \frac{LSE - LIE}{6\sigma}, \quad (6.1)$$

y compara el rango de valores permitido al proceso con el rango natural del mismo. En el caso en el que no se conoce el valor de σ , se estimará. Entonces,

$$C_p = \frac{LSE - LIE}{6\hat{\sigma}}. \quad (6.2)$$

Veamos a continuación cómo se interpreta el índice de capacidad.

- Si $C_p = 1$, $LSE - LIE = 6\hat{\sigma}$; en este caso, la amplitud natural del proceso y el rango de valores que le permitimos son iguales: se dice que el proceso es marginalmente capaz de cumplir con los límites de especificación.
- Si $C_p < 1$, $LSE - LIE < 6\hat{\sigma}$; en este caso, el rango permitido al proceso es menor que la amplitud del mismo: habrá, por tanto, elementos disconformes con los límites de especificación, o lo que es lo mismo, el proceso no es capaz de cumplir con los límites de especificación.
- Si $C_p > 1$, $LSE - LIE > 6\hat{\sigma}$; el rango de valores permitido al proceso es más amplio que el rango de valores que está tomando el proceso. En este caso, se dice que el proceso es capaz de cumplir con los límites de especificación.

Cuanto mayor sea C_p mayor será la capacidad del proceso de cumplir con las especificaciones. En particular, se considera que el nivel mínimo de calidad es $C_p \geq 1.33$, y uno de los objetivos más extendidos hoy en día en las compañías es obtener un valor de $C_p \geq 1.6$.

En las Figuras 6.1, 6.2 y 6.3 tenemos los tres casos posibles.

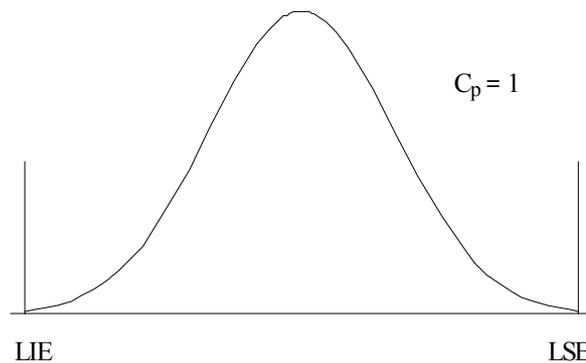


Figura 6.1: Proceso marginalmente capaz

El índice de capacidad C_p es adimensional por lo que permite comparar la capacidad de procesos diferentes.

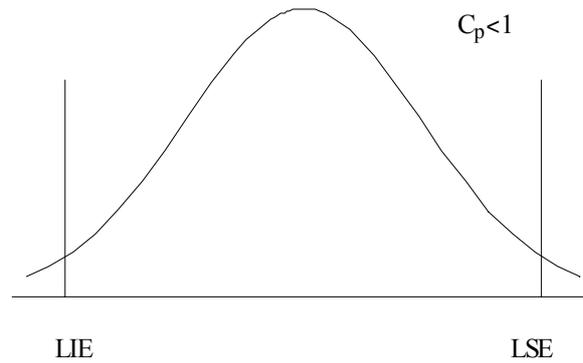


Figura 6.2: Proceso no capaz

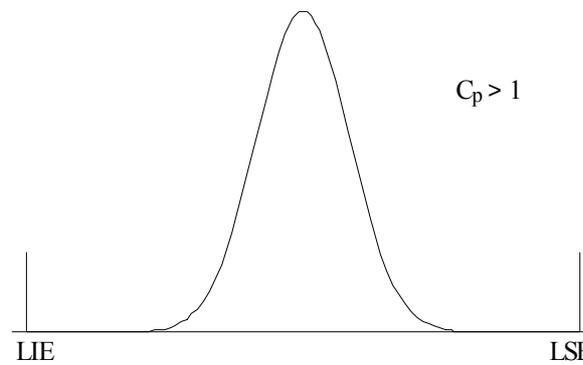


Figura 6.3: Proceso capaz

Ejemplo 71 *Comparamos la capacidad de los siguientes dos procesos:*

1. *Las estimaciones de la media y la desviación típica para este proceso son $\hat{\mu} = 14.0018$, $\hat{\sigma} = 0.00091$, respectivamente, y los límites de especificación son $LIE = 13.995$ y $LSE = 14.005$.*
2. *Las estimaciones de la media y la desviación típica para este proceso son $\hat{\mu} = 14.0018$, $\hat{\sigma} = 0.168$, respectivamente, y los límites de especificación son $LIE = 13$ y $LSE = 14$.*

Para el primer proceso

$$C_p = \frac{14.005 - 13.995}{6 \cdot 0.00091} = 1.8315$$

para el segundo proceso

$$C_p = \frac{14 - 13}{6 \cdot 0.168} = 0.99206$$

El primer proceso tiene un $C_p > 1.6$, luego tiene una alta capacidad de cumplir con los límites de especificación aunque estos sean muy estrechos. Sin embargo, a pesar de que para el segundo proceso la amplitud de los límites de especificación es mucho mayor, su $C_p < 1$ nos indica que el proceso no es capaz de cumplir con las especificaciones. La diferencia está en la desviación típica, como vemos, uno de los objetivos primordiales ha de ser el de reducir la variabilidad.

6.5.1 Índice de Funcionamiento

El índice C_p nos da una medida de la capacidad que tiene el proceso para cumplir las especificaciones, pero tiene el inconveniente de que no tiene en cuenta la localización del proceso, por lo que, si el proceso no está centrado en el valor nominal de la media puede ocurrir que, aunque la amplitud natural del proceso sea pequeña, se estén produciendo más elementos que no cumplen con las especificaciones de los que se debiesen esperar. Por ejemplo, si dos procesos tienen el mismo C_p y uno de ellos no está centrado, tendríamos la situación que se presenta en las Figuras 6.4 y 6.5.

En los dos casos el proceso es *potencialmente* capaz de cumplir con los límites de especificación, pero en el segundo caso, al estar desviada la media del proceso hacia un valor mayor que el nominal, se produce una mayor proporción de elementos disconformes.

Para evitar este problema se define un nuevo índice para medir la capacidad del proceso, el *índice de funcionamiento* C_{pk} ,

$$C_{pk} = \min \left[\frac{LSE - \hat{\mu}}{3\hat{\sigma}}, \frac{\hat{\mu} - LIE}{3\hat{\sigma}} \right], \quad (6.3)$$

donde $\hat{\mu}$ y $\hat{\sigma}$ son las estimaciones de μ y σ respectivamente.

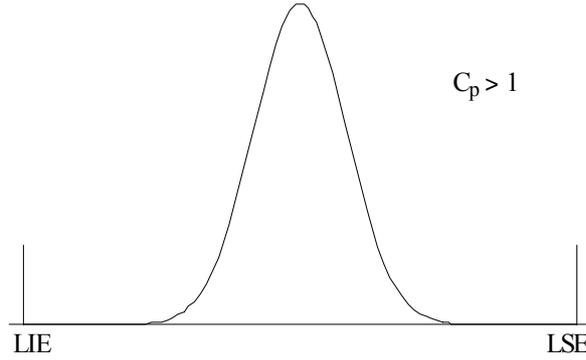


Figura 6.4: Proceso capaz y centrado

El subíndice k hace referencia al llamado k -factor:

$$k = \frac{\left| \frac{LSE+LIE}{2} - \hat{\mu} \right|}{\frac{LSE-LIE}{2}}, \quad (6.4)$$

y mide la longitud de cuánto se aleja la media del valor central entre los límites de especificación (que es el valor objetivo). Se puede demostrar que $0 \leq k \leq 1$, de modo que cuanto más cercano esté k a 0, más centrado estará el proceso en el valor objetivo para la media.

Simples operaciones algebraicas relacionan los índices C_p y C_{pk} de la siguiente manera

$$C_{pk} = (1 - k) C_p. \quad (6.5)$$

Esta relación nos dice que $C_{pk} \leq C_p$ y se alcanza la igualdad cuando el proceso está centrado en el valor objetivo para la media; cuanto más cercanos estén estos dos valores entre sí, más centrado estará el proceso.

Con estos dos índices obtenemos una imagen bastante clara de cómo funciona un proceso en relación a los límites de especificación.

Ejemplo 72 Calculemos el C_p y el C_{pk} para un proceso con estimaciones de la media y de la desviación típica (obtenidas de los datos una vez que el proceso está bajo control estadístico) $\hat{\mu} = 3.9911$ y $\hat{\sigma} = 0.00924$, y con

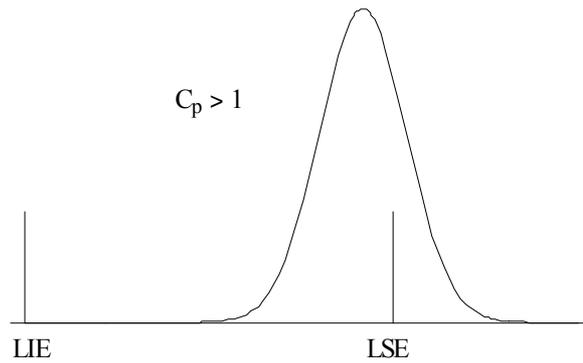


Figura 6.5: Proceso capaz no centrado

límites de especificación $LIE = 3.95$ y $LSE = 4.05$:

$$k = \frac{\left| \frac{4.05+3.95}{2} - 3.99911 \right|}{\frac{4.05-3.95}{2}} = 0.0178$$

$$C_p = \frac{4.05 - 3.95}{6 \cdot 0.00924} = 1.8038$$

$$C_{pk} = (1 - 0.0178)(1.8038) = 1.7717$$

Observamos que $C_p > 1.6$, por lo que el proceso es altamente capaz de cumplir las especificaciones. Además, $k = 0.0178$, por tanto podemos deducir que el proceso está centrado en el valor objetivo. Esto último también puede deducirse de la proximidad de C_p y C_{pk} .

6.5.2 Índices de Capacidad Unilaterales

Hasta ahora hemos visto dos índices, C_p y C_{pk} , para evaluar la capacidad de un proceso con límites de especificación superior e inferior. En muchos procesos la característica de calidad que nos interesa sólo tiene límites de especificación a un lado. Para ellos es interesante disponer de índices de capacidad unilaterales. Estos índices siguen la misma filosofía que el índice de capacidad, pero están diseñados para la mitad del rango de valores del proceso que nos interesa. Más específicamente:

Para un proceso con límite de especificación inferior se define el *índice*

de capacidad inferior, C_{pl} , por

$$C_{pl} = \frac{\hat{\mu} - LIE}{3\hat{\sigma}}. \quad (6.6)$$

Análogamente, para un proceso con límite de especificación superior se define el *índice de capacidad superior*, C_{pu} por

$$C_{pu} = \frac{LSE - \hat{\mu}}{3\hat{\sigma}}. \quad (6.7)$$

Es evidente que

$$C_{pk} = \min [C_{pl}, C_{pu}].$$

Además, en distribuciones normalmente distribuidas, estos índices están relacionados con la proporción de elementos fabricados fuera de los límites de especificación, concretamente

$$\begin{aligned} P(X > LSE) &= P\left(Z > \frac{LSE - \hat{\mu}}{\hat{\sigma}}\right) = P(Z > 3C_{pu}) \\ P(X < LIE) &= P\left(Z < \frac{LIE - \hat{\mu}}{\hat{\sigma}}\right) = P(Z < -3C_{pl}) \end{aligned} \quad (6.8)$$

donde $Z \rightsquigarrow N(0, 1)$. Con (6.8) podemos convertir fácilmente los índices de capacidad unilaterales en proporciones de elementos disconformes.

6.5.3 Índice de Capacidad Recíproco C_r

Este índice es el inverso de C_p , es decir, $C_r = 1/C_p$. Por lo tanto, su interpretación es la siguiente:

- Si $C_r > 1$, el proceso no es capaz.
- Si $C_r < 0.752$, el proceso es capaz.
- Si $0.752 \leq C_r \leq 1$, el proceso es marginalmente capaz, es decir, el proceso es capaz pero precisa de un control estricto.

El índice C_r mide la cantidad de intervalo de especificación, $LSE - LIE$, que utiliza el proceso. Cuanto mayor sea la variabilidad natural del proceso, $6\hat{\sigma}$, más utilizará éste el intervalo de especificación y, recíprocamente,

procesos centrados con altos C_p tienden a utilizar menos amplitud de la tolerancia.

Cuando C_r se expresa como un porcentaje se conoce con el nombre de *porcentaje de especificación* utilizado por el proceso. Nótese que el índice C_{pk} no se utiliza en esta definición. La razón es que normalmente se quiere saber cuánto se utiliza del rango de tolerancia en las circunstancias más ideales (i.e. cuando el proceso está centrado) y es el índice C_p (no C_{pk}) el que resume el potencial del proceso para cumplir las especificaciones.

Ejemplo 73 Consideremos un proceso con $C_p = 0.61$. Sabemos que un $C_p < 1$ indica un proceso que no tiene capacidad para cumplir las especificaciones. Para este proceso el índice de capacidad recíproco es,

$$\frac{1}{0.61} = 1.64.$$

Esto es, el proceso utiliza el 164% del rango de tolerancia. Por tanto, el proceso requiere más tolerancia (64% más) de la que se permite, aunque el proceso esté perfectamente centrado en el valor objetivo. Esto es porque la variación del proceso es demasiado grande para cumplir con las especificaciones.

6.5.4 Índice de Capacidad Modificado C_{pm}

Este índice es similar al C_p , salvo que utiliza el valor nominal μ_N en lugar de la estimación de la media (\bar{x}) en el cálculo de la desviación típica. Por tanto su definición es

$$C_{pm} = \frac{LSE - LIE}{6\hat{\sigma}_m},$$

donde la estimación de la desviación típica del proceso, $\hat{\sigma}_m$, (a partir de k muestras de tamaño n) es, en este caso,

$$\hat{\sigma}_m = \frac{\sum_{j=1}^k s_{jm}}{k},$$

$$s_{jm} = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \mu_N)^2}{n-1}}.$$

Este índice es más conservador que C_p , es decir, suele salir más bajo. Es más recomendable cuando se poseen pocos datos.

6.5.5 Coeficiente n -sigma

Este coeficiente consiste en la expresión de los límites de especificación en función de la desviación típica del proceso de la forma $LSE-LIE = 2n\sigma$. Esta definición de los límites de especificación equivale a $LSE = \mu + n\sigma$, y $LIE = \mu - n\sigma$. Por tanto, un índice de capacidad $C_p = 1$ equivale a 3-sigma y $C_p = 2$ equivale a un nivel de calidad 6-sigma (six sigma). Actualmente se considera que una organización ha alcanzado una cota muy alta de calidad si su nivel es 6-sigma. Por esta razón, a los programas de formación en técnicas estadísticas para la calidad se les suele denominar *programas 6-sigma*.

Es frecuente interpretar el nivel de calidad n -sigma en términos de número de artículos defectuosos por millón. Para ello hay dos posibles alternativas que suelen producir confusión. La primera alternativa consiste en emplear estrictamente su significado (siempre se supone normalidad). Por ejemplo, si el proceso está en estado de control y el nivel de calidad es 3-sigma, el 0.27% de los artículos son defectuosos. También puede decirse en este caso que 2700 artículos por millón serán defectuosos.

Una segunda interpretación, popularizada por la empresa *Motorola*, consiste en calcular el número de defectos por millón cuando se produce un desajuste en la media de $\pm 1.5\sigma$, considerándose que cualquier causa asignable puede fácilmente provocar este nivel de desajuste.

Ejemplo 74 *Calculemos la proporción, en tantos por millón, de artículos defectuosos que se producen con y sin desajuste de 1.5σ , para un coeficiente de 1-sigma.*

Que el coeficiente sea de 1-sigma significa que $LSE = \mu + \sigma$, $LIE = \mu - \sigma$.

- Sin desajuste,

$$\begin{aligned}
 P[X > LSE] + P[X < LIE] &= P[X > \mu + \sigma] + P[X < \mu - \sigma] = \\
 &= P\left[Z > \frac{\mu + \sigma - \mu}{\sigma}\right] + P\left[Z < \frac{\mu - \sigma - \mu}{\sigma}\right] = \\
 &= P[Z > 1] + P[Z < -1] = \\
 &= 2P[Z < -1] = 0.317310 \Rightarrow 317310 \text{ por millón}
 \end{aligned}$$

donde $Z \rightsquigarrow N(0, 1)$.

- Con desajuste de $+1.5\sigma$, es decir, la media del proceso es $\mu + 1.5\sigma$

$$\begin{aligned}
 P[X > LSE] + P[X < LIE] &= P[X > \mu + \sigma] + P[X < \mu - \sigma] = \\
 &= P\left[Z > \frac{\mu + \sigma - \mu - 1.5\sigma}{\sigma}\right] + P\left[Z < \frac{\mu - \sigma - \mu - 1.5\sigma}{\sigma}\right] = \\
 &= P[Z > -0.5] + P[Z < -2.5] = 0.697670 \Rightarrow 317310 \text{ por millón}
 \end{aligned}$$

donde $Z \rightsquigarrow N(0, 1)$.

En la Tabla 6.1 se muestra la proporción (en tantos por millón) de artículos defectuosos que se producen con y sin el desajuste de 1.5σ , en función del coeficiente $n - \text{sigma}$ y suponiendo normalidad.

$n - \text{sigma}$	En control	Desajuste de $\pm 1.5\sigma$
1	317310	697670
2	45500	308769
3	2700	66810
4	64	6210
5	0.57	233
6	0.002	3.4

Tabla 6.1: Proporción de artículos defectuosos por millón

En la Tabla 6.1 podemos observar que un sistema con nivel de calidad de 6-sigma y una desviación del objetivo de $\pm 1.5\sigma$ producirá, en término medio, 3.4 artículos defectuosos por millón.

Ejemplo 75 *Un artículo defectuoso entre 160 artículos equivale a una proporción de*

$$\frac{1}{160} = 0.00625 = 6250 \times 10^{-6} \Rightarrow 6250 \text{ por millón}$$

que es, aproximadamente, un nivel de calidad de 4-sigma (con un desajuste de $\pm 1.5\sigma$) y equivale a un C_p de

$$C_p = \frac{8\sigma}{6\sigma} = 1.33$$

que es considerado el nivel mínimo de calidad que se debe tener.

6.6 Errores de Medida e Índices de Capacidad

En el cálculo de los índices de capacidad no hemos tenido en cuenta los errores de medida que se producen al tomar las observaciones, o lo que es lo mismo, hemos supuesto que los métodos para obtener las observaciones tienen la precisión suficiente como para que el error de la medida fuese despreciable respecto de la medida en sí. Sin embargo esta situación no es muy realista, la variación del error influye en la variación de las observaciones obtenidas y, por tanto, tendremos que incluir el error en los estudios de capacidad. Esto es lo que se pretende en esta sección.

En primer lugar, observamos que cualquier observación del proceso x_m es el resultado de la suma de dos factores, el verdadero valor de la característica x y el error de medida producido ϵ , es decir

$$x_m = x + \epsilon. \quad (6.9)$$

De (6.6) el único valor que nosotros podemos obtener es x_m , ni x , ni ϵ los podremos conocer en realidad. Sin embargo, podremos estimar la magnitud de ϵ y, a partir de ahí, separar la variación del proceso de la variación del error.

Si consideramos que el error de medida es estadísticamente independiente del valor observado x_m , la variación de las observaciones obtenidas σ_m^2 se puede expresar como

$$\sigma_m^2 = \sigma^2 + \sigma_\epsilon^2, \quad (6.10)$$

donde σ_ϵ^2 denota la varianza de los errores de medida y σ^2 es la verdadera varianza del proceso. Por lo tanto, podemos combinar el conocimiento del tamaño de σ_ϵ^2 con σ_m^2 para estimar la verdadera variabilidad del proceso σ . Recordemos que σ_m es la estimación utilizada en los índices de capacidad, no la verdadera (y desconocida) σ .

Una forma de medir el efecto de los errores de medida es estimar el porcentaje del intervalo de especificación que utiliza el término error, ϵ . Suponiendo que dichos errores de medida siguen una distribución normal, podemos utilizar el índice C_r del error (que vamos a notar C_r^ϵ) para medir dicho porcentaje, esto es

$$C_r^\epsilon = \frac{6\sigma_\epsilon}{LSE - LIE}.$$

Despejamos σ_ϵ :

$$\sigma_\epsilon = \frac{C_r^\epsilon (LSE - LIE)}{6}$$

y sustituimos en (6.10):

$$\sigma_m^2 = \sigma^2 + \left[\frac{C_r^\epsilon (LSE - LIE)}{6} \right]^2.$$

Ahora, si llamamos C_p^* al índice de capacidad obtenido a partir de las observaciones, podemos escribir:

$$C_p^* = \frac{LSE - LIE}{6\sigma_m} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + \left[\frac{C_r^\epsilon (LSE - LIE)}{6} \right]^2}}.$$

Simplificando se tiene:

$$\begin{aligned} C_p^* &= \frac{1}{\frac{6}{LSE - LIE} \sqrt{\sigma^2 + \left[\frac{C_r^\epsilon (LSE - LIE)}{6} \right]^2}} = \frac{1}{\sqrt{\left(\frac{6\sigma}{LSE - LIE} \right)^2 + (C_r^\epsilon)^2}} = \\ &= \frac{1}{\sqrt{\left(\frac{1}{C_p} \right)^2 + (C_r^\epsilon)^2}}, \end{aligned} \tag{6.11}$$

donde C_p denota el “verdadero” índice de capacidad, basado en la (desconocida) variación del proceso, σ .

Si no hay variación debida al error, es decir, no hay errores de medida, $\sigma_\epsilon = 0$, entonces $C_r^\epsilon = 0$ y $C_p^* = C_p$.

Además, si hacemos $C_p \rightarrow \infty$, esto es $\sigma \rightarrow 0$ (reducimos la variabilidad del proceso), entonces el índice de capacidad obtenido a partir de las observaciones, C_p^* , está acotado superiormente:

$$C_p^* \leq \frac{1}{C_r^\epsilon}.$$

Por lo tanto, la magnitud de los errores de medida impone una cota superior para C_p^* . Por otra parte, despejando en (6.11); se tiene:

$$C_p = \frac{1}{\sqrt{\frac{1}{(C_p^*)^2} - (C_r^\epsilon)^2}},$$

ecuación que puede utilizarse para estimar la verdadera capacidad del proceso después de haber calculado el error de las mediciones.

Ejemplo 76 *Supongamos un proceso químico con límites de especificación 100 ± 20 y donde las mediciones están sujetas a una variación de $\sigma_\epsilon = 6$. Por tanto:*

$$C_r^\epsilon = \frac{6\sigma_\epsilon}{LSE - LIE} = 0.9.$$

Esto es, los errores de medida utilizan un 90% del intervalo de especificación. Por consiguiente, una cota superior del índice de funcionamiento de las observaciones es:

$$C_p^* \leq \frac{1}{C_r^\epsilon} = 1.1,$$

y esto aunque la variación del proceso pudiese reducirse a 0. Entonces, los esfuerzos para obtener un índice de funcionamiento superior serían inútiles.

6.7 Capacidad del Proceso: Atributos

En las secciones previas, se ha juzgado la capacidad del proceso para el caso en que la característica de calidad observada era de tipo variable. En esta sección vamos a introducir brevemente las nociones básicas de capacidad de un proceso en el que la característica de calidad observada sea del tipo atributo, específicamente el diagrama de la proporción de elementos

disconformes o bien el número de disconformidades por unidad de inspección.¹ Como en el caso de las variables, medir la capacidad del proceso tiene sentido cuando éste está bajo control.

6.7.1 Proporción de Disconformes

Cuando se está controlando un proceso vigilando la fracción disconforme, p , la medida del rendimiento del proceso se define como la proporción de elementos no conformes en todo el periodo de evaluación. Esto es igual al valor \bar{p} que se calcula como valor de la línea central del diagrama de p . Así, si se han tomado m muestras de tamaño n cada una,

$$\bar{p} = \frac{\sum_{i=1}^m n\hat{p}_i}{nm}.$$

Pues bien, en este caso, la capacidad del proceso se mide en términos de elementos conformes,

$$1 - \bar{p},$$

ó $100(1 - \bar{p})\%$.

En el caso de atributos no hay un valor estándar de $1 - \bar{p}$ que indique cómo de buena es la capacidad del proceso. ? se recomienda como una capacidad aceptable un 98-99% para imperfecciones no muy importantes, y una capacidad de 99.9% ó más en otro caso.

6.7.2 Número de Disconformidades

En este caso los datos necesarios para medir la capacidad de un proceso se obtienen del correspondiente diagrama de control, el diagrama de c o el diagrama de u .

El rendimiento del proceso, en este caso se mide mediante la razón de ocurrencias de disconformidades por unidad, es decir, corresponde a la \bar{u} , línea central del diagrama de u , esto es, si obtenemos m muestras de tamaño

¹Nótese que un elemento se clasifica como conforme o no conforme al compararlo con las especificaciones

n , y x_i es el número de defectos observados en la muestra i ,

$$\bar{u} = \frac{\sum_{i=1}^m x_i}{nm}.$$

Esta medida del rendimiento del proceso también se suele expresar en términos de defectos por 100 unidades, $100\bar{u}$, e incluso en defectos por millón.

Análogamente al caso de la proporción de elementos disconformes, no hay un valor establecido para el cuál se considere que el rendimiento del proceso es positivo.

6.8 Ejercicios

1. En un proceso de fabricación de fuentes de energía eléctrica la variable de calidad de interés es el voltaje de salida. Se mide dicha variable en grupos de tamaño 5, y se obtuvieron los siguientes datos:

N ^o de muestra	\bar{x}_i	r_i
1	103	4
2	102	5
3	104	2
4	105	11
5	104	4
6	106	3
7	102	7
8	105	2
9	106	4
10	104	3

- (a) ¿Cuáles son los límites de control en el periodo de vigilancia? ¿y en el periodo base?
- (b) ¿Cuáles son los límites de tolerancia natural del proceso?
- (c) ¿Cuál sería la proporción de artículos disconformes si las especificaciones de calidad señalan que el voltaje de salida ha de encontrarse entre 99 y 107?.
- (d) ¿Es el proceso capaz? ¿Está centrado?
2. Se toman muestras de un proceso de manufactura a intervalos regulares. Se mide cierta característica de calidad, distribuida normalmente, y se calculan los valores de \bar{X} y S para cada muestra ($n = 4$). Después de 30 muestras, se tiene que

$$\sum_{i=1}^{30} \bar{x}_i = 1560, \quad \sum_{i=1}^{30} s_i = 42$$

- (a) Calcular los límites de control para los diagramas de control de \bar{X} y S .
- (b) Suponga que todos los puntos de ambos diagramas caen entre los límites de control. ¿Cuáles son los límites de tolerancia natural del proceso?

- (c) Si los límites de especificación son 50 ± 7 , ¿qué puede concluirse acerca de la capacidad del proceso de producir artículos conformes con las especificaciones?
- (d) A la vista de los resultados obtenidos en el apartado anterior, ¿qué sugerencias harías para mejorar la capacidad del proceso?
3. Las siguientes cifras son las medias y rangos de muestras de tamaño 5. Los datos corresponden a la profundidad del resalte de las cabezas de las bombas de fragmentación producidas durante la guerra por la American Stove Company. Las mediciones están en pulgadas:

Nº de muestra	\bar{x}	r	Nº de muestra	\bar{x}	r
1	0.4402	0.015	16	0.4362	0.015
2	0.4390	0.018	17	0.4380	0.019
3	0.4448	0.018	18	0.4350	0.008
4	0.4432	0.006	19	0.4378	0.011
5	0.4428	0.008	20	0.4384	0.009
6	0.4382	0.010	21	0.4392	0.006
7	0.4358	0.011	22	0.4378	0.008
8	0.4440	0.019	23	0.4362	0.016
9	0.4366	0.010	24	0.4348	0.009
10	0.4368	0.011	25	0.4338	0.005
11	0.4360	0.011	26	0.4366	0.014
12	0.4402	0.007	27	0.4346	0.009
13	0.4332	0.008	28	0.4374	0.015
14	0.4356	0.017	29	0.4339	0.024
15	0.4314	0.010	30	0.4368	0.014

- (a) Con base a las 20 primeras muestras establezca un diagrama de \bar{X} y otro de R .

$$\sum_{i=0}^{20} \bar{x}_i = 8.7632 \quad \sum_{i=0}^{20} r_i = 0.238$$

¿Está el proceso bajo control estadístico durante el periodo de las muestras 21-30?

- (b) Los límites de especificación de la American Stove Company son 0.445 ± 0.015 pulgadas. Analizar la capacidad del proceso

para cumplir con las especificaciones a partir de las estimaciones obtenidas en el periodo base. Interpretar los resultados

- (c) ¿Qué proporción de elementos no cumple con las especificaciones?. ¿Y si la media del proceso se desplaza a 0.4315?

4. Considera los dos procesos siguientes (el tamaño muestral $n = 5$)

Proceso A	Proceso B
$\bar{x}_A = 100$	$\bar{x}_B = 105$
$\bar{S}_A = 3$	$\bar{S}_B = 1$

Las especificaciones son 100 ± 10 . Calcular todos los índices de capacidad e interpretarlos. ¿Qué proceso prefieres?

5. Una característica de calidad normalmente distribuida se vigila mediante los diagramas de \bar{X} y R . Estos diagramas tienen los siguientes parámetros ($n = 4$):

\bar{X}	R
LSC = 626	LSC = 18.795
LC = 620	LC = 8.236
LIC = 614	LC = 0

ambos diagramas muestran control estadístico.

- (a) ¿Cuál es la desviación típica estimada del proceso?
- (b) Supongamos que un diagrama de S va a sustituir al diagrama de R . ¿Cuáles serán los parámetros del diagrama de S ?
- (c) Si las especificaciones del producto son 610 ± 15 , ¿cuál es la estimación de la proporción de elementos no conformes que produce el proceso?
- (d) ¿Qué se puede hacer para reducir la proporción de elementos disconformes?
6. Los siguientes datos son la medias muestrales y los rangos para 24 muestras de tamaño 5 tomadas de un proceso de producción de arandelas. Las medidas se hacen sobre el diámetro interior de las arandelas, y sólo se recogen los tres últimos decimales.

- (a) Establecer los diagramas de control para este proceso.
- (b) Si las especificaciones de dicho diámetro son 0.5030 ± 0.0010 , ¿cuál es el porcentaje de elementos no conformes producidos por el proceso?
- (c) Calcula e interpreta los índices de capacidad del proceso suponiendo normalidad en los datos.

Muestra	\bar{x}_i	R_i	Muestra	\bar{x}_i	R_i
1	34.5	3	13	35.4	8
2	34.2	4	14	34.0	6
3	31.6	4	15	37.1	5
4	31.5	4	16	34.9	7
5	35.0	5	17	33.5	4
6	34.1	6	18	31.7	3
7	32.6	4	19	34.0	8
8	33.8	3	20	35.1	4
9	34.8	7	21	33.7	2
10	33.6	8	22	32.8	1
11	31.9	3	23	33.5	3
12	38.6	9	24	34.2	2

7. Para un cierto proceso, los diagramas de control de \bar{X} y R se basan en muestras de tamaño 5 y tienen líneas centrales de 14.5 y 1.163 respectivamente. Si el proceso tiene límites de especificación 12 y 16, calcular C_p , C_{pk} , C_{pl} y C_{pu} .
8. Para vigilar un proceso se toman muestras de tamaño 5 y se ha calculado el recorrido medio de las mismas. El valor hallado ha sido $R = 0.10$. Las tolerancias de las piezas objeto de control son: Tolerancia máxima: 10.84, Tolerancia mínima: 10.02. ¿Podría decir si el proceso tiene capacidad suficiente?.

Capítulo 7

Introducción al Muestreo para la Aceptación

7.1 Introducción

El muestreo para la aceptación es otra herramienta estadística para asegurar la calidad de un producto. Es similar a otras técnicas estadísticas donde se hace inferencia sobre el proceso basándose en muestras de elementos del proceso.

Históricamente, el muestreo para la aceptación se ha aplicado tanto al final de la producción como para comprobar la calidad de las materias primas que entran en la empresa. A este muestreo en particular se le conoce como *control de recepción* y ha estado motivado porque durante mucho tiempo, la filosofía subyacente en las relaciones cliente-proveedor ha sido la desconfianza. En la actualidad compradores y proveedores han asumido la necesidad de que la relación entre ellos se base en el respeto mutuo. Para ello muchas empresas se someten por parte de una tercera entidad a la evaluación de una certificación tipo ISO 9000, de modo que si una empresa posee una de estas certificaciones, sus clientes demuestran una confianza que elimina la necesidad de las inspecciones de los lotes recibidos. Sin embargo, en determinadas ocasiones es necesario realizar un muestreo para la aceptación, como por ejemplo, por un cambio de proveedor, o bien por la sospecha de que el proveedor, incluso certificado, ha disminuido la calidad de la materia prima suministrada.

Por otra parte, el muestreo para la aceptación en las fases finales de la

producción es una de las primeras actividades de una empresa que quiere comenzar a establecer un sistema de calidad. Sin embargo tiene el inconveniente de que no es preventivo, cuando se detecta un producto defectuoso ya es demasiado tarde. El movimiento de la inspección del producto al control de procesos es fundamental en la filosofía del Control de Calidad moderno. Está ampliamente reconocido que se debe introducir el control de procesos tan pronto como sea posible, realizándose la inspección en las fases tempranas de la producción, de forma que las técnicas de inspección den paso a las técnicas de prevención.

Los mecanismos del muestreo para la aceptación son muy simples. Ya sea al final de la producción, o con materias primas que se reciben en la compañía, los productos están agrupados en lotes, y la idea general es decidir qué hacer con el lote, aceptarlo o rechazarlo en su totalidad, a partir de la inspección de una parte de los artículos del lote.

El muestreo para la aceptación puede dividirse en dos grandes grupos:

- Muestreo por atributos: en la inspección los artículos se dividen en defectuosos o no defectuosos según cumplan con los requerimientos. En este caso se rechaza el lote si el número de elementos no conformes en la muestra excede un número fijado previamente.
- Muestreo por variables: en la inspección se mide una variable cuantitativa. Se rechaza el lote siempre que la media muestral de la muestra observada exceda algún valor especificado.

Por su relativa simplicidad, los planes de muestreo para la aceptación basados en datos de atributos son utilizados con más frecuencia que los basados en datos de variables.

Los fundamentos del muestreo para la aceptación se desarrollaron durante 1925-1927 en los laboratorios Bell (igual que los diagramas de control). Después de esto, el muestreo para la aceptación se aplicó esporádicamente hasta la 2ª Guerra Mundial, cuando se incorporó a los estándares militares. El uso de estos planes en los trabajos militares y de defensa ayudó a promover el amplio uso del muestreo para la aceptación, que continúa hasta nuestros días. Uno de estos planes, *MIL-STD-105E*, es muy utilizado.

En la actualidad, el muestreo para la aceptación se utiliza en las situaciones en las que no se pueden aplicar los diagramas de control o en los que

todavía hay causas especiales. Algunos ejemplos en los que está justificado el muestreo para la aceptación son los siguientes:

1. Para una inspección de la materia prima.
2. Para una primera inspección del artículo.
3. Para procesos que todavía no están bajo control.
4. Para aquellos casos en que la prueba es destructiva.
5. Cuando no es posible una inspección al 100%, ya sea porque el coste económico es muy elevado, o no es tecnológicamente factible, o cuando se necesitaría tanto tiempo que se vería afectada la producción.

7.2 Formación de los lotes

Los planes más usuales para el muestreo para la aceptación se basan en el muestreo de lotes. Llamaremos *lote* a un conjunto de artículos, en general numeroso, del que tenemos que decidir si lo adquirimos o lo rechazamos en función de la proporción de elementos defectuosos que contenga. Llamaremos *muestra* al conjunto de artículos que extraemos del lote y que serán los únicos que examinaremos para ver si son defectuosos o aceptables.

Entre los distintos tipos de lotes que se pueden formar (lotes de manufactura, lotes de envío) los que se utilizan para el muestreo para la aceptación son los *lotes de inspección*. Muchas veces, los lotes de inspección se predeterminan por la forma en la que se seleccionan los productos o por la forma en que se envían. Otras veces, se puede determinar el tamaño y la composición del lote de inspección. Cuando esto es posible hay que seguir dos principios básicos para formar los lotes:

1. Los lotes deben ser homogéneos (interiormente).
2. Si son homogéneos, son mejores los lotes grandes que los pequeños.

La primera razón para querer lotes grandes es que admiten tamaños muestrales grandes, con lo que la decisión de aceptar o rechazar el lote es más fiable. Pero sólo se puede permitir un tamaño de lote grande si éste es homogéneo. La homogeneidad dentro del lote significa que los elementos

del lote deben ser lo más similares posible en su fecha y en su forma de producción. Lotes mixtos, que incluyen unidades de distintos niveles de calidad, pueden llevar a que los lotes que pasan un plan de muestreo tengan una menor calidad que la inicialmente prevista.

7.3 Planes de muestreo único para atributos

7.3.1 Introducción

El muestreo basado en lotes con datos de atributos es la forma más simple y común del muestreo para la aceptación. Un *plan de muestreo para la aceptación de atributos* se caracteriza por un tamaño muestral, n , y un número de aceptación, c , (c es un entero no negativo). Se toma una muestra de tamaño n de un lote dado, después se compara el número de elementos no conformes en la muestra con el número de aceptación y se acepta o se rechaza el lote completo. Específicamente, si $X =$ “número de elementos no conformes en la muestra”, entonces: Si $X > c$ se rechaza el lote y si $X \leq c$ se acepta el lote.¹

Ejemplo 77 *Supongamos que se desea tomar una decisión sobre un lote de 1500 artículos, y que el plan de muestreo establecido es que $n = 50$, $c = 1$. En este caso se procederá como sigue: se toma una muestra de tamaño 50, se cuentan el número de elementos defectuosos x , si $x \leq 1$ se acepta el lote, en otro caso se rechaza.*

Por lo tanto se desea decidir si el lote es bueno o no en función de n y c , es decir, establecer un plan de muestreo consiste en determinar n y c de forma que se cumplan unas condiciones establecidas.

7.3.2 Curva Característica Operativa

En la práctica nunca se sabe la calidad de un lote por adelantado. Si un lote está compuesto enteramente por elementos conformes, entonces un plan de muestreo aceptará siempre dicho lote (si suponemos que no hay errores de clasificación), ya que $X = 0$. Igualmente, lotes que contengan

¹Nótese que el tamaño del lote N , no aparece como parte del plan de muestreo

únicamente elementos defectuosos se rechazarán. Sin embargo, la mayoría de los lotes que se inspeccionan están entre esos dos extremos y la proporción de no conformes, p , en cada uno de dichos lotes es desconocida y varía de lote a lote. Para solucionar este problema, los planes de muestreo se evalúan mediante el cálculo de la probabilidad de aceptación, $P_a = P[X \leq c]$, sobre el rango de los posibles valores de p . Representando P_a frente a p ($0 \leq p \leq 1$), se puede ver cómo reaccionará el plan frente a cualquier situación (esto es, a cualquier nivel de elementos no conformes en el lote). Por esta razón, dicho gráfico se llama *curva característica de operación* o *curva característica operativa* (curva OC) del plan en particular. Esto es, la curva característica de operación es un gráfico que expresa, para un plan de muestreo concreto, la probabilidad de aceptar un lote en función del porcentaje p de artículos defectuosos existentes en el lote. Para representar esta probabilidad utilizaremos la notación $OC(p) = P_a$, por ejemplo, tal y como hemos comentado antes si $p = 0$ aceptaremos siempre el lote, $OC(0) = 1$ (salvo errores de clasificación), por el contrario, $OC(1) = 0$. Como se puede esperar, P_a no crece al aumentar p , por lo tanto las curvas OC generalmente decrecen cuando p aumenta de 0 a 1. Véase por ejemplo la Figura 7.1.

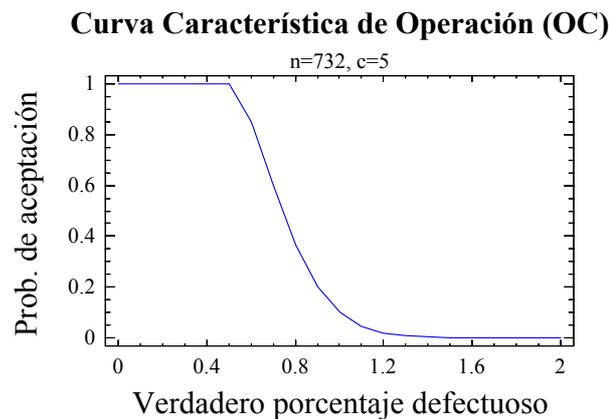


Figura 7.1: Curva OC de un plan de muestreo

Al observar la curva OC para un plan de muestreo en particular, se debe decidir si éste ofrece el grado de control deseado sobre el lote. Por ejemplo,

un plan de muestreo con una P_a grande es inadecuado para lotes de baja calidad. Por otro lado, es posible que un plan de muestreo que dé una alta protección contra la aceptación de lotes de baja calidad, rechace un número elevado de lotes de alta calidad. Por tanto, en un muestreo para la aceptación siempre existe el riesgo de tomar una decisión equivocada. Las equivocaciones pueden ser en dos sentidos:

- Error Tipo I: rechazar un lote bueno.
- Error Tipo II: aceptar un lote malo.

La efectividad de un plan de muestreo se determinará por la capacidad de minimizar estos dos errores.

En el caso del muestreo para la aceptación la probabilidad de error Tipo I y la probabilidad de error Tipo II, reciben unos nombres particulares:

- *Riesgo del productor*. Es la probabilidad de que un lote de alta calidad sea rechazado, es decir, la probabilidad de error Tipo I, α .
- *Riesgo del consumidor*. Es la probabilidad de que un lote de baja calidad sea rechazado, es decir, la probabilidad de error Tipo II, β .²

La decisión de los valores de p que corresponden a un lote de alta calidad o de baja calidad es subjetiva. Habitualmente se fijan previamente entre cliente y proveedor un *nivel de calidad aceptable* (AQL , *Acceptance Quality Level*) y un *nivel de calidad rechazable* (LQL *Limiting Quality Level* o $LTPD$ *Lot Tolerance Percentage Defective*) de forma que P_a sea alta cuando $p = AQL$ y P_a sea baja cuando $p = LTPD$.

Con la notación anteriormente introducida, el riesgo del productor y el riesgo del consumidor, serán $\alpha = 1 - OC(AQL)$ y $\beta = OC(LTPD)$, respectivamente.

En la práctica, los planes de muestreo se determinan fijando a priori los valores de α , AQL , β , $LTPD$. Habitualmente, los valores de α están alrededor del 5% y β alrededor del 10%. Por ejemplo, si $AQL = 0.02$ y $LTPD = 0.1$ entonces, en general, habrá que determinar n y c de forma

²Nótese que la calidad de un lote se expresa en valores de p , lotes con valores pequeños de p son lotes de buena calidad, mientras que lotes con valores elevados de p son lotes de baja calidad

que P_a no debe exceder 0.95 para lotes con $p = AQL$ y P_a no debe exceder 0.01 para lotes con $p = 0.10$. En esencia, este procedimiento establece dos puntos en la curva OC, se debe escoger el plan (seleccionar n y c) cuya curva OC pasa lo más cerca posible de estos dos puntos.

7.3.3 Probabilidad de aceptación de un lote

Para calcular P_a lo primero que hay que determinar es la manera en la que se va a utilizar el plan de muestreo. Esencialmente hay dos formas distintas de utilizar un plan de muestreo: como un método para juzgar un lote individual o como un método para juzgar una corriente continua de lotes. A los planes de muestreo para juzgar lotes individuales los llamaremos *Planes Tipo A* y los que juzgan series de lotes *Planes Tipo B*. Los más usuales son los de Tipo B. Conocer si el plan es de Tipo A o de Tipo B es esencial para calcular correctamente P_a y, por lo tanto, para encontrar la curva OC del plan.

El cálculo de probabilidades para los planes de Tipo A utilizan la distribución Hipergeométrica. Ya que sólo se analiza un lote de tamaño N , sólo hay $N+1$ posibles valores de p para cada P_a que se puede calcular. Esto es, puesto que sólo puede haber $0, 1, 2, \dots, N$ elementos no conformes en el lote, los únicos posibles valores para p son $0, \frac{1}{N}, \frac{2}{N}, \dots, 1$. Para cada valor de p , la composición del lote (el número de elementos conformes y el número de elementos no conformes) es conocida. Por lo tanto, $P_a = P[X \leq c]$ se puede calcular mediante la distribución Hipergeométrica. Nótese que la curva OC debe ser escalonada ya que p nunca toma valores distintos de $0, \frac{1}{N}, \frac{2}{N}, \dots, 1$.

Ejemplo 78 *Para inspeccionar un lote de tamaño $N = 10$ se utiliza un plan de muestreo $n = 8, c = 1$. Puesto que hay 20 elementos en el lote, P_a sólo se puede calcular en los puntos, $0, 0.05, 0.1, \dots, 0.9, 1$. Así, $P_a = P[X = 0] + P[X = 1]$, probabilidades que se calculan sabiendo que $X \rightsquigarrow H(20, 8, p.20)$. Dichas probabilidades se recogen en la Tabla 7.1.*

Por las restricciones impuestas por el pequeño tamaño del lote $N = 20$, $P_a = 0$ para valores $p \geq 0.7$. La curva OC para este plan se muestra en la Figura 7.2

Si por el contrario el plan de muestreo se utiliza para el seguimiento continuo de lotes, utilizaremos un plan Tipo B y su curva OC asociada se

p	$P[X \leq 1]$	p	$P[X \leq 1]$
0	1	0.55	$\frac{27}{8398}$
0.05	1	0.60	$\frac{97}{125970}$
0.10	$\frac{81}{95}$	0.65	$\frac{1}{9690}$
0.15	$\frac{187}{285}$	0.70	0
0.20	$\frac{451}{969}$	0.75	0
0.25	$\frac{99}{323}$	0.80	0
0.30	$\frac{121}{646}$	0.85	0
0.35	$\frac{341}{3230}$	0.90	0
0.40	$\frac{2277}{41990}$	0.95	0
0.45	$\frac{11}{442}$	1	0
0.50	$\frac{83}{8398}$		

Tabla 7.1: Probabilidades de aceptación para los distintos valores de p

calcula con la distribución Binomial.

Ejemplo 79 Si el plan $n = 8$, $c = 1$, del Ejemplo 78, fuese utilizado para un plan Tipo B, para cada p , $P_a = P[X \leq 1] = P[X = 0] + P[X = 1] = (1 - p)^8 + \binom{8}{1}p(1 - p)^7 = (1 - p)^8 + 8p(1 - p)^7$. En la Figura 7.3 se muestra la curva OC para este plan. En la Figura 7.4 se comparan ambas curvas OC.

En los ejemplos considerados, las curvas difieren algo ya que para el plan Tipo A una muestra de tamaño $n = 8$ es relativamente grande comparada con el tamaño del lote $N = 20$. Para planes en los que el tamaño del lote es mucho más grande que el tamaño muestral ($N < 0.1n$) ocurren dos cosas: la curva OC se dibuja en muchos más puntos ($N + 1$ para ser exactos), dando una apariencia más suave, y las probabilidades de los planes Tipo A y Tipo B llegan a ser casi indistinguibles, ya que la distribución hipergeométrica se aproxima a la Binomial cuando $N < 0.1n$.

También se puede aplicar muestreo para la aceptación cuando se están contando disconformidades por elemento en lugar de elementos no conformes. En estas aplicaciones lo habitual es expresar el número de disconformidades encontrado, tanto como una proporción por elemento inspeccionado, como una proporción por 100 elementos inspeccionados. El procedimiento es análogo al de contar elementos no conformes. Para calcular la curva OC para estos planes se utiliza la distribución de Poisson.

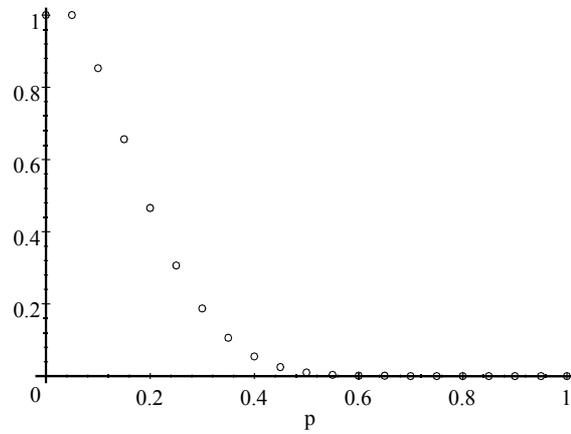


Figura 7.2: Curva OC de un plan de muestreo Tipo A, $N = 20$, $n = 8$, $c = 1$

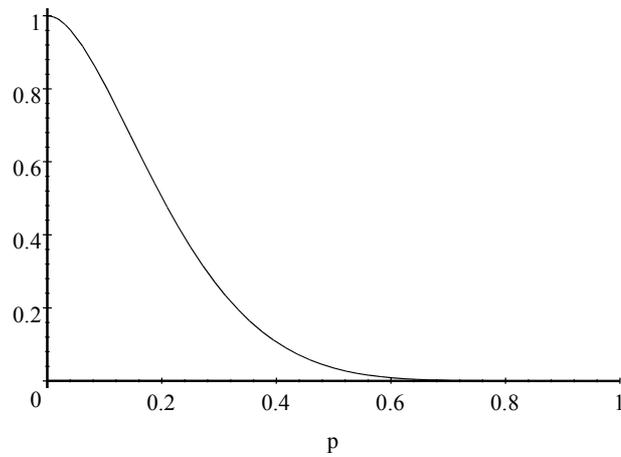


Figura 7.3: Curva OC de un plan de muestreo Tipo B, $n = 8$, $c = 1$

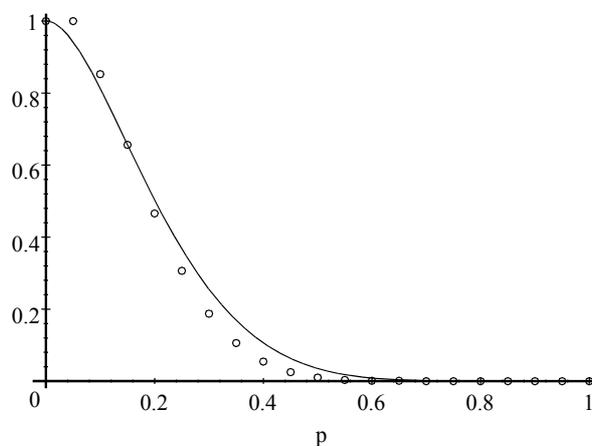


Figura 7.4: Comparación de las dos curvas OC anteriores

Ahora c se refiere al número total de no conformidades permitidas en los n elementos inspeccionados. Entonces los valores de p (interpretada como “errores por elemento”) primero se convierten en la correspondiente razón por n elementos, y después se calculan las probabilidades de Poisson.

Ejemplo 80 Para un plan de muestreo $n = 8$, $c = 2$, un valor de $p = 0.1$ no conformidades por elemento, en primer lugar se traslada a una razón de 0.8 no conformidades por 8 elementos. Por lo tanto:

$$\begin{aligned}
 P_a &= P[X \leq 2] = P[X = 0] + P[X = 1] + P[X = 2] = \\
 &= \frac{e^{-\lambda} \lambda^0}{0!} + \frac{e^{-\lambda} \lambda^1}{1!} + \frac{e^{-\lambda} \lambda^2}{2!} = \\
 &= e^{-\lambda} \left(1 + \lambda + \frac{\lambda^2}{2} \right) = e^{-0.8} \left(1 + 0.8 + \frac{0.8^2}{2} \right) = 0.9526
 \end{aligned}$$

De la misma forma se calculan P_a para distintos valores de p para obtener la curva OC del plan.

7.3.4 Determinación de un Plan de Muestreo

Como hemos comentado anteriormente, determinar un plan de muestreo único para atributos consiste en determinar el tamaño de la muestra, n , y el número de aceptación, c . También hemos comentado que, en la práctica, se fijan unos valores a priori de α , AQL , β , $LTPD$ y se determina el plan de muestreo que verifique $\alpha = 1 - OC(AQL)$ y $\beta = OC(LTPD)$. Esto puede ser imposible, y aunque podamos dibujar varias curvas OC y seleccionar el plan de muestreo cuya curva más se aproxime a dichos puntos, esta solución es muy ineficiente. Sin embargo, se puede encontrar alguna aproximación razonable si el tamaño muestral es suficientemente grande, ya que entonces se puede utilizar la aproximación a la Normal. Veamos un ejemplo.

Ejemplo 81 *Se desea diseñar un plan de muestreo simple con los siguientes parámetros: $AQL = 0.03$, $\alpha = 0.05$, $LTPD = 0.07$, $\beta = 0.05$. Vamos a suponer que el tamaño muestral n , que es uno de los dos parámetros que hemos de determinar, será lo suficientemente grande como para poder utilizar la aproximación a la Normal. Sea X el número de artículos defectuosos en la muestra. Entonces:*

$$X \rightsquigarrow B(n, p) \approx N\left(np, \sqrt{np(1-p)}\right)$$

Los valores de $AQL = 0.03$ y $\alpha = 0.05$ significan que la probabilidad de rechazar un lote con una proporción de defectuosos igual a 0.03 es 0.05, esto es:

$$P(X > c \mid p = 0.03) = 0.05$$

Aplicando la aproximación de la Binomial por la Normal y tipificando:

$$P(X > c) = P\left(Z > \frac{c - n0.03}{\sqrt{n0.03(1-0.03)}}\right) = 0.05$$

De donde, consultando las tablas de la Normal:

$$\frac{c - n0.03}{\sqrt{n0.03(1-0.03)}} = 1.64$$

Despejando c :

$$c = 0.03n + 0.28\sqrt{n}$$

Análogamente, de $LTPD = 0.07$, $\beta = 0.05$.

$$P(X \leq c | p = 0.07) = 0.05$$

Mediante la aproximación Normal y tipificando:

$$P(X \leq c) = P\left(Z \leq \frac{c - n0.07}{\sqrt{n0.07(1 - 0.07)}}\right) = 0.05$$

de donde

$$\frac{c - n0.07}{\sqrt{n0.07(1 - 0.07)}} = -1.64$$

y despejando c ,

$$c = 0.07 - 0.42\sqrt{n}$$

y así obtenemos un sistema de dos ecuaciones y dos incógnitas cuya solución es:

$$n \simeq 307$$

$$c \simeq 14$$

Por tanto, el plan que se busca es el siguiente: se toma una muestra de tamaño $n = 307$. Se rechaza el lote si se encuentran $c = 14$ artículos defectuosos o más en dicha muestra.

7.3.5 Sistemas de muestreo para la aceptación

Aunque las curvas OC son relativamente fáciles de calcular, pueden ser muy ineficaces si se tienen que repetir estos cálculos cada vez que se necesita seleccionar un plan de muestreo. Más aún, probablemente sería necesario calcular las curvas OC para varios planes, ya que es difícil estimar los valores exactos de n y c necesarios para alcanzar niveles específicos de riesgo del productor y de riesgo del cliente. Para evitar este problema, y en un esfuerzo para hacer los planes de muestreo aún más versátiles, se han agrupado un gran número de planes de muestreo con propiedades similares en *sistemas de muestreo* cuyas curvas OC se publican e incluyen como una parte del sistema. El *MIL-STD-105E* es uno de estos sistemas de muestreo.

Seleccionar el plan apropiado de uno de los sistemas de muestreo se reduce a elegir entre las curvas OC del sistema, la que dé el nivel de protección deseado.

Hay varias formas de agrupar planes de muestreo para formar un sistema. Por ejemplo, se podría elaborar un sistema donde los planes de muestreo estuviesen organizados según riesgos específicos de productores y consumidores. Veamos a continuación algunas de las medidas utilizadas para formar los sistemas de muestreo. Algunas han sido introducidas previamente de una forma breve.

Nivel de Calidad Aceptable, AQL

Como hemos señalado anteriormente el AQL (*Acceptance Quality Level*) especifica un valor particular del nivel de calidad del productor.

Para utilizar un sistema AQL , primero se especifica un valor para el AQL (por ejemplo $AQL = 0.01$) y después se utilizan las tablas del sistema publicado para seleccionar valores de n y c que puedan alcanzar ese AQL . Esto significa, implícitamente, que P_a debe ser alta cuando $p = AQL$. Como consecuencia, el sistema de muestreo debe especificar el valor (o rango de valores) que se utilizan para P_a cuando $p = AQL$. Lo más común es elegir $P_a = 0.95$ a $p = AQL$, aunque se puede especificar cualquier valor de P_a si así se desea. El sistema *MIL-STD-105E* está organizado según distintos niveles de AQL .

Nivel de Calidad Rechazable, $LTPD$

Esta medida también fue introducida con anterioridad. Se comentó que el *nivel de calidad aceptable* (en inglés se suele denominar *lot tolerance poercentage defective, LTPD*), es la proporción de artículos defectuosos en un lote que es inaceptable para el comprador. Por convenio, la notación $LTPD$ usualmente se refiere al punto en la curva OC en el que $P_a = 0.10$. Por lo tanto, para asegurar que la mayoría de los lotes que contienen un 20% de elementos defectuosos ($p = 0.20$) son rechazados por un plan de muestreo, se ha de especificar un $LTPD$ de 0.20 (equivalentemente $LTPD = 20\%$). Un plan de muestreo con este $LTPD$ tendrá una curva OC para la cual $P_a = 0.10$ para lotes con $p = 0.20$; es decir, el 90% de estos lotes (en los que $p = 0.20$) son rechazados por el plan. Por supuesto, la probabilidad de aceptación se puede fijar en cualquier valor, no solamente en $P_a = 0.10$,

pero la mayoría de los sistemas *LTPD* se basan en una $P_a = 0.10$. Un sistema de muestreo basado en *LTPD* es el sistema de *Dodge-Romig* (?).

Límite de la Calidad Media de Salida, AOQL

Los sistemas de muestreo *AOQL* se basan en el *límite de la calidad media de salida* (*Average Outgoing Quality Limit*). Esta medida de rendimiento se aplica sólo a ciertas formas de muestreo de inspección en los cuáles los lotes rechazados se estudian al 100% y cualquier elemento no conforme que se encuentra se reemplaza por uno bueno. Esta forma de muestreo y subsecuente corrección de los lotes rechazados se llama *inspección de rectificación* y se explicará con más detalle en la siguiente sección.

Bajo la inspección de rectificación, hay una cota superior del porcentaje medio de no conformes que será dirigida al cliente. Es fácil entender porqué existe dicha cota: la mayoría de los lotes de alta calidad son aceptados (por lo tanto el porcentaje de no conformes en esos lotes es pequeño) y la mayoría de los lotes de baja calidad serán rechazados y rectificados (por tanto, el porcentaje de no conformes en estos lotes debe estar próximo a 0). Así representando la *calidad media de salida* (*AOQ*, *Average Outgoing Quality*), es decir, el porcentaje de elementos defectuosos que hay en los lotes, frente al porcentaje de no-conformes, p , en los lotes suministrados para la inspección resulta el gráfico de la Figura 7.5.

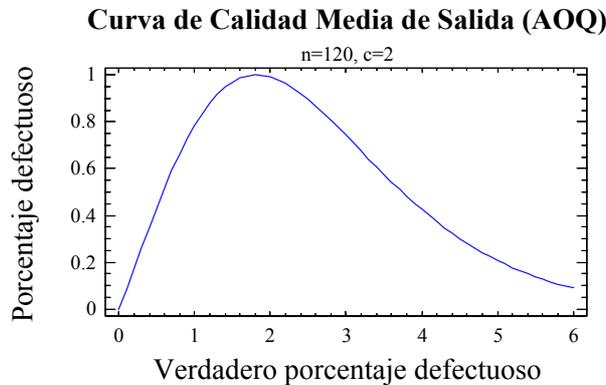


Figura 7.5: Gráfico AOQ para un plan de muestreo

El punto máximo de esa curva se llama *límite de la calidad de salida AOQL*. El *AOQL* representa el peor porcentaje posible de no conformes que se puede esperar en los lotes vigilados bajo un plan de muestreo particular. Tanto el plan *MIL-STD-105E* como el de la *Dodge-Romig* incluyen tablas de *AOQL* y gráficas para ayudar al usuario a elegir el plan correcto cuando se está utilizando inspección de rectificación.

Inspección Total Media, ATI

La *inspección total media (ATI, Average Total Inspection)* es el número medio de elementos inspeccionados. Esta medida también tiene interés únicamente en el caso de realizar muestreo de rectificación. Cada lote rechazado añade $N - n$ elementos adicionales para ser inspeccionados. Por supuesto, cuanto peor sea la calidad del lote (i.e. p aumenta), más y más lotes se rechazan y aumenta la calidad total de inspección. En general, la inspección total media, *ATI*, se puede representar frente a p . Estos gráficos se utilizan para determinar el costo adicional asociado con la inspección de rectificación. Un ejemplo se muestra en la Figura 7.6.

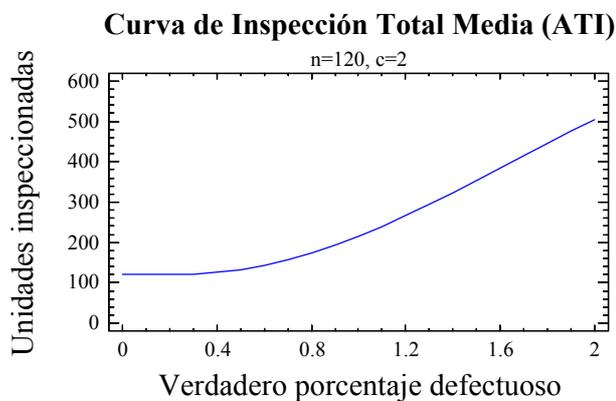


Figura 7.6: Gráfica de la Inspección Total Media de un plan de muestreo

7.4 Inspección de rectificación

Como hemos comentado brevemente en la sección anterior, la inspección de rectificación es una forma de muestreo para la aceptación en la cual cualquier elemento no conforme que se encuentre en los lotes rechazados se reemplaza antes de que esos lotes se envíen a los clientes. En la inspección de rectificación tiene sentido definir (para cualquier proporción de no conformes p) la calidad media de salida (AOQ) que es la proporción media de elementos no conformes en los lotes enviados. Mas aún, en el caso del muestreo de inspección, es habitual que cualquier elemento no conforme en una muestra de tamaño n de un lote de tamaño N también sea reemplazado por otro bueno antes de tomar la decisión de si se rechaza o se acepta el lote. Por lo tanto la proporción de elementos no conformes en lotes aceptados en una inspección de rectificación es:

$$p \frac{(N - n)}{N}$$

para cualquier valor particular de p . En la Figura 7.7 se resume esta idea.

Con ello, la calidad media de los lotes aceptados (AOQ) mediante muestreo de rectificación es:

$$AOQ = p \left(\frac{N - n}{N} \right) P_a + 0$$

Cuando el tamaño de la muestra es pequeño comparado con el tamaño del lote, el factor $\frac{N-n}{N}$ estará próximo a 1, por lo que una buena aproximación del AOQ estará dada por:

$$AOQ \simeq p \cdot P_a$$

El punto p en el cual la curva AOQ alcanza el máximo se puede calcular

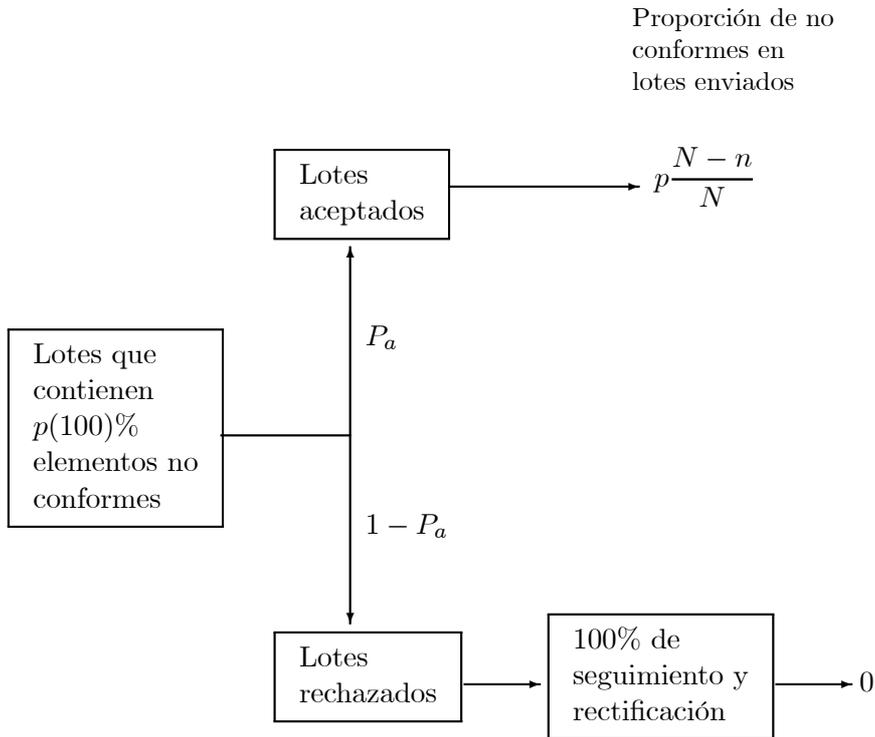


Figura 7.7: Inspección de rectificación

mediante cálculo elemental:

$$\begin{aligned}
 AOQ &= p \frac{N-n}{N} P_a, \quad c = \text{n}^\circ \text{ de aceptación} \quad P_a = P(X \leq c) \\
 \frac{\partial AOQ}{\partial p} &= \frac{\partial}{\partial p} \left[p \frac{N-n}{N} P_a \right] = \frac{N-n}{N} \frac{\partial}{\partial p} [p P(X \leq c)] = \\
 &= \frac{N-n}{N} \frac{\partial}{\partial p} \left[p \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k} \right] = \\
 &= \frac{N-n}{N} \frac{\partial}{\partial p} \left[\sum_{k=0}^c \binom{n}{k} p^{k+1} (1-p)^{n-k} \right] = \\
 &= \frac{N-n}{N} \sum_{k=0}^c \binom{n}{k} \left[p^{k+1} (n-k) p^{n-k-1} (-1) + (1-p)^{n-k} (k+1) p^k \right] = \\
 &= \frac{N-n}{N} \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k-1} [(k+1) - (n+1)p]
 \end{aligned}$$

Igualamos a 0:

$$\begin{aligned} \frac{N-n}{N} \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k-1} [(k+1) - (n+1)p] &= 0 \\ \sum_{k=0}^c \binom{n}{k} p^k (1-p)^{n-k-1} [(k+1) - (n+1)p] &= 0 \end{aligned} \quad (7.1)$$

Resolviendo esa ecuación, obviando las raíces triviales $p = 1$ o $p = 0$, se obtiene el valor deseado p_{\max} .

Recordemos que el valor de AOQ cuando $p = p_{\max}$ es límite de la calidad media de salida o $AOQL$. Hay que tener cuidado con la interpretación del $AOQL$ ya que no es la máxima proporción de no conformes que puede ser enviada bajo un plan en particular. Es posible que algunos lotes enviados puedan tener niveles de calidad un poco peores que el $AOQL$. Sin embargo, si promediamos sobre todos los lotes donde $p = p_{\max}$, el $AOQL$ es el máximo que ese promedio puede tomar.

Ejemplo 82 *Se desea calcular el $AOQL$ asociado a un plan $n = 25$, $c = 1$ Sustituyendo en (7.1)*

$$\begin{aligned} \sum_{k=0}^1 \binom{25}{k} p^k (1-p)^{25-k-1} [(k+1) - 26p] \\ = \binom{25}{0} p^0 (1-p)^{25-1} (1 - 26p) + \binom{25}{1} p (1-p)^{25-2} (2 - 26p) \\ = (1-p)^{23} [(1-p)(1-26p) + 25p(2-26p)] = 0 \end{aligned}$$

p_{\max} es una de las raíces de:

$$\begin{aligned} (1-p)(1-26p) + 25p(2-26p) &= 0 \\ -624p^2 + 23p + 1 &= 0 \end{aligned}$$

Las raíces son $p = 0.0625$ y $p = -0.0256!!$, por lo tanto $p_{\max} = 0.0625$.

Para calcular $AOQL$ sustituimos p_{\max} :

$$\begin{aligned} AOQL &= p_{\max} \left(\frac{N-n}{N} \right) P_a = p_{\max} P(X \leq 1) \left(\frac{N-n}{N} \right) = \\ &= (0.0625) \left[\binom{25}{0} 0.0625^0 (1 - 0.0625)^{25} + \binom{25}{1} 0.0625 (1 - 0.0625)^{24} \right] \left(\frac{N-n}{N} \right) = \\ &= 0.0332 \left(\frac{N-n}{N} \right) \end{aligned}$$

Si suponemos que N es mucho mayor que $n \Rightarrow AOQL \simeq 0.0332$ ó 3.32%

Tal y como se comentó en la sección anterior, en el muestreo de rectificación otra medida de interés es la inspección total media. Cada lote inspeccionado requiere que se inspeccione una muestra de tamaño n y, además, cada lote rechazado contribuye con $N - n$ elementos adicionales. Ya que la probabilidad de rechazar un lote es $1 - P_a$, el número medio de elementos inspeccionados por lote, o el promedio total de inspección, ATI , por lote se puede estimar por:

$$ATI = n + (N - n)(1 - P_a)$$

7.5 Muestreo doble y múltiple

7.5.1 Introducción

Una generalización de los planes de muestreo único son los planes de muestreo múltiple, que incluyen dos o más pasos de muestreo. En cada paso la decisión es aceptar, rechazar o pasar al siguiente paso de muestreo y tomar muestras adicionales antes de decidir sobre el lote. Con este procedimiento los lotes de baja calidad se rechazarán después de la primera muestra, los lotes de alta calidad se aceptarán después de la primera muestra, y los restantes lotes requerirán más inspección antes de tomar una decisión sobre ellos.

En particular, el plan de muestreo doble es una extensión inmediata del único. En primer lugar, se extrae una muestra de tamaño n_1 y se cuenta el número de defectuosos, X_1 . Si este número es muy elevado, digamos mayor o igual a cierto valor r_1 , el lote se rechaza. Si X_1 es muy bajo, menor o igual que un valor c_1 el lote se acepta. Sin embargo, si X_1 está entre estos

dos valores se concluye que la muestra no arroja evidencia suficiente para tomar una decisión. En ese caso, se toma una segunda muestra de tamaño n_2 del lote y se evalúa el número de defectuosos en dicha muestra, X_2 . Si $X_1 + X_2$ es mayor que cierta cantidad r_2 el lote se rechaza definitivamente. En caso contrario se acepta.

Un plan de muestreo doble se especifica mediante la elección de cinco números: los tamaños muestrales n_1 y n_2 para las dos etapas de muestreo, junto con los respectivos números de aceptación y rechazo (c_1 , c_2 , r_1 , r_2) para cada paso. Aunque parece que hay seis parámetros, sólo se escogen cinco, ya que el número de rechazo en el segundo paso está exactamente determinado por c_2 . Esto es, si denotamos por X_1 el número de elementos no conformes en la primera muestra y, en caso de que sea necesaria una segunda muestra, X_2 es el número de elementos no conformes que contiene esta muestra, el número total de no conformes encontrados en ambas muestras se denota $Y = X_1 + X_2$. Utilizamos la misma notación que para planes de muestreo único, por lo tanto c_1 y c_2 representan los números de aceptación para las dos fases. Los números de rechazo r_1 y r_2 también son necesarios en cada paso. La operación de un plan de muestreo es, por tanto:

Paso 1 Tomar una muestra de tamaño n_1 del lote.

Si $X_1 \leq c_1 \Rightarrow$ aceptar el lote.

Si $X_1 \geq r_1 \Rightarrow$ rechazar el lote.

Si $c_1 < X_1 < r_1 \Rightarrow$ ir al Paso 2.

Paso 2 Tomar una muestra adicional de tamaño n_2 del lote, contar el número de no-conformes X_2 y encontrar el número total de no-conformes en las dos muestras $Y = X_1 + X_2$.

Si $Y \leq c_2 \Rightarrow$ aceptar el lote.

Si $Y \geq r_2 \Rightarrow$ rechazar el lote ($r_2 = c_2 + 1$).

Ya que la decisión se debe tomar en o antes de la segunda muestra, el número de rechazo r_2 ha de ser $c_2 + 1$. La otra restricción natural que se debe cumplir para que el plan sea factible es $c_1 \leq r_1 \leq c_2$. En la Figura 7.8 se muestra un esquema de este plan de muestreo.

La extensión al caso general de muestreo múltiple es fácil. Utilizando la misma notación, elegimos los números de aceptación y de rechazo c_i y r_i para cada paso i . X_i denota el número de elementos no conformes en

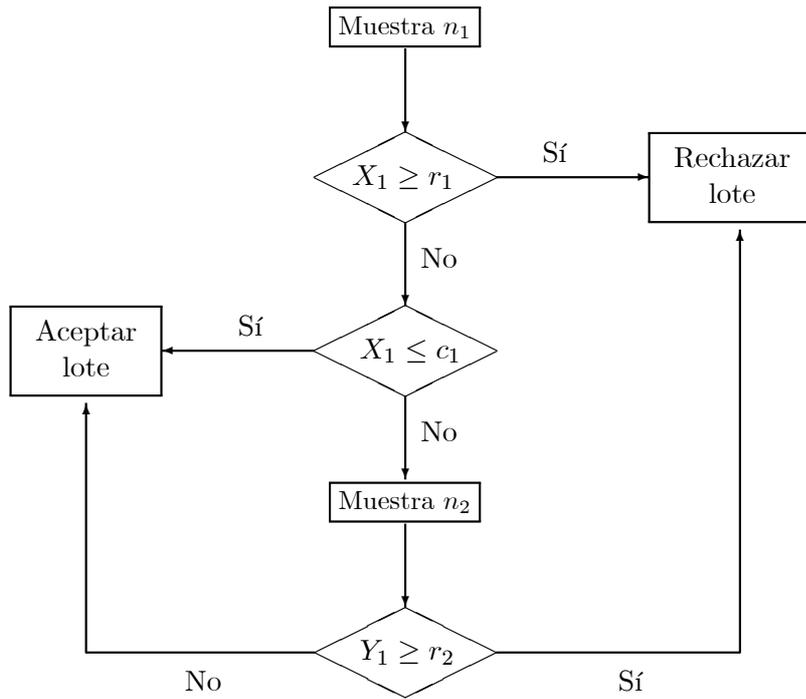


Figura 7.8: Esquema de un plan de muestreo doble

la i -ésima muestra. Entonces el número acumulado en el paso i -ésimo, $Y_i = X_1 + X_2 + \dots + X_i$ se utiliza para decidir si aceptar o rechazar el lote en el i -ésimo paso:

Paso 1 Tomar una muestra de tamaño n_1 del lote y sea X_1 el número de no conformes encontrados.

Si $X_1 \leq c_1 \Rightarrow$ aceptar el lote.

Si $X_1 \geq r_1 \Rightarrow$ rechazar el lote.

Si $c_1 < X_1 < r_1 \Rightarrow$ ir al paso siguiente.

Paso 2 Tomar una muestra adicional de n_2 del lote, contar el número de no-conformes X_2 , y formar el conteo acumulado $Y_2 = X_1 + X_2$.

Si $Y_2 \leq c_2 \Rightarrow$ aceptar el lote.

Si $Y_2 \geq r_2 \Rightarrow$ rechazar el lote.

Si $c_2 < Y_2 < r_2 \Rightarrow$ ir al paso siguiente.

⋮

Paso k Tomar una muestra adicional de tamaño n_k , contar el número de elementos no conformes X_k , y formar el conteo acumulativo $Y_k = X_1 + X_2 + \dots + X_k$.

Si $Y_k \leq c_k \Rightarrow$ aceptar el lote.

Si $Y_k \geq r_k \Rightarrow$ rechazar el lote. ($r_k = c_k + 1$).

Como en el muestreo doble, no se permiten más de k pasos por lo tanto tenemos la restricción $r_k = c_k + 1$.

Los planes de muestreo doble y múltiple también pueden ser recogidos en sistemas de muestreo, lo mismo que los planes de muestreo con una única muestra, mediante la medida de alguna medida de rendimiento. El *MIL-STD-105E*, por ejemplo, incluye la opción de seleccionar muestreos únicos, dobles o múltiples con el *AQL*.

7.5.2 Curvas OC para muestreo doble y múltiple

El resultado final de un plan de muestreo múltiple es el mismo que el de cualquier plan de muestreo para la aceptación: un lote entero se acepta

o se rechaza. Por lo tanto, la probabilidad de aceptar un lote, P_a , se puede calcular y representar frente a la calidad del lote, p , para formar la curva característica de operación asociada con el plan, curva OC. Sin embargo, con muestreo múltiple, los cálculos son un poco más complicados, ya que la aceptación de un lote (o el rechazo) puede ocurrir en cualquiera de los k pasos. En esta sección ilustraremos estos cálculos para las curvas OC de los planes de muestreo doble. Las curvas OC para planes de muestreo múltiple son más complejas pero se hacen esencialmente de la misma forma.

Los parámetros que describen un plan de muestreo doble son (n_1, c_1, r_1) y $(n_2, c_2, r_2 = c_2 + 1)$.

Un lote se acepta cuando $X_1 \leq c_1$ o (si se necesita una segunda muestra) cuando $Y = X_1 + X_2 \leq c_2$. Estos dos sucesos son mutuamente excluyentes con lo que P_a es la suma de esas dos probabilidades. La primera de esas componentes, $P(X_1 \leq c_1)$ se calcula mediante una Binomial. La segunda componente $P(Y \leq c_2)$ se calcula mediante probabilidades condicionadas:

$$P(Y \leq c_2) = P(Y \leq c_2 \mid c_1 < X_1 < r_1) P(c_1 < X_1 < r_1)$$

Para simplificar las expresiones utilizamos la notación $b(s, n)$ para denotar la probabilidad Binomial de encontrar s elementos no conformes en una muestra de tamaño n , esto es, $b(s, n) = \binom{n}{s} p^s (1-p)^{n-s}$. Entonces, para cada valor de p , P_a se puede escribir como:

$$\begin{aligned} P_a &= P(X_1 \leq c_1) + P(Y \leq c_2) = \\ &= P(X_1 \leq c_1) + P(Y \leq c_2 \mid c_1 < X_1 < r_1) P(c_1 < X_1 < r_1) \\ &= P(X_1 \leq c_1) + \\ &P(X_2 \leq c_2 - X_1 \mid c_1 + 1 \leq X_1 \leq r_1 - 1) P(c_1 + 1 \leq X_1 \leq r_1 - 1) \\ &= \sum_{s=0}^{c_1} b(s, n_1) + \sum_{s=c_1+1}^{r_1-1} \left[b(s, n_1) \sum_{j=0}^{c_2-s} b(j, n_2) \right] \end{aligned}$$

Ejemplo 83 Supongamos que un plan de muestreo doble con parámetros $(n_1 = 10, c_1 = 0, r_1 = 2)$ y $(n_2 = 12, c_2 = 1, r_2 = c_2 + 1 = 2)$ se considera como posible reemplazamiento de un plan de muestreo único $(n = 13, c = 1)$. Para comparar los dos planes estudiamos las respectivas curvas OC:

Para el plan de muestreo simple, para cada p ,

$$P_a = \sum_{s=0}^1 b(s, 13) = (1-p)^{13} + \binom{13}{1} p (1-p)^{12} = (1-p)^{12} (1 + 12p)$$

Para el muestreo doble:

$$\begin{aligned}
 P_a &= \sum_{s=0}^0 b(s, 10) + \sum_{s=1}^1 \left[b(s, 10) \sum_{j=0}^{1-s} b(j, 12) \right] = \\
 &= b(0, 10) + b(1, 10) b(0, 12) = \\
 &= \binom{10}{0} (1-p)^{10} + \binom{10}{1} p (1-p)^9 \binom{12}{0} (1-p)^{12} = \\
 &= (1-p)^{10} [1 + 10p(1-p)^{11}]
 \end{aligned}$$

Para comparar los dos planes dibujamos las dos curvas OC (Figura 7.9)
 Las curvas OC están muy próximas para $0 \leq p \leq 0.05$. Para valores

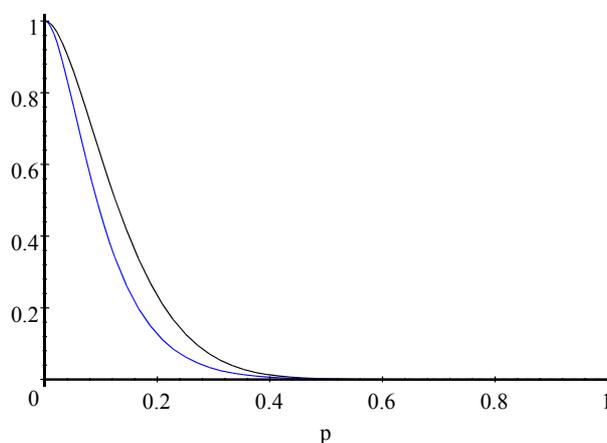


Figura 7.9: Curvas OC de un plan de muestreo simple (línea negra) y un plan de muestreo doble (línea azul)

mayores de $p = 0.05$, la curva OC para el muestreo único es más alta que para el doble, hasta $p = 0.5$, donde las curvas vuelven a juntarse. Concluimos que para AQL menores que 0.05, el plan de muestreo doble proporciona esencialmente el mismo grado de protección que el muestreo único y proporciona una mayor protección contra los lotes de baja calidad (i.e. cuando $p > 0.05$).

Si se supone que el proveedor es capaz de enviar lotes de calidad $p \leq 0.05$ la mayor parte del tiempo, entonces el plan de muestreo doble reducirá el total de número de elementos inspeccionados (ya que una muestra de tamaño 10 usualmente es suficiente para sentenciar dichos lotes).

Para lotes de menor calidad, no está claro si el plan de muestreo doble reduce o incrementa el trabajo de inspección.

7.5.3 Tamaño Muestral Medio

En muestreo doble y múltiple el tamaño muestral que se requiere para alcanzar la decisión de aceptar o rechazar el lote varía. Algunas veces la decisión se puede alcanzar en la primera muestra de tamaño n_1 , otras veces es necesario un mayor muestreo. Ya que la cantidad de inspección que se requiere es siempre un factor importante a considerar, encontrar el tamaño muestral medio, *ASN* (*Average Sample Number*) es interesante para planes de muestreo múltiple.

Para un plan de muestreo doble con parámetros (n_1, c_1, r_1) y $(n_2, c_2, r_2 = c_2 + 1)$, el tamaño muestral mínimo es n_1 , ya que todos los lotes han de pasar por lo menos por el primer paso. Como una muestra adicional de tamaño n_2 se necesita únicamente si $c_1 < X_1 < r_1$, el tamaño muestral medio será:

$$ASN = n_1 + n_2 P(c_1 < X_1 < r_1) = n_1 + n_2 \sum_{s=c_1+1}^{r_1-1} b(s, n_1)$$

Representar la curva de *ASN* frente a p ayuda a determinar la cantidad de esfuerzo de inspección que uno puede esperar cuando se utiliza el plan.

Ejemplo 84 En el ejemplo anterior, el plan de muestreo doble ($n_1 = 10$, $c_1 = 0$, $r_1 = 2$), ($n_2 = 12$, $c_2 = 1$, $r_2 = c_2 + 1 = 2$) se comparó con el plan de muestreo único ($n = 13$, $c = 1$). Calculemos el *ASN* del plan de muestreo doble para los posibles valores de p junto con el del plan de muestreo único. Ya que para el plan de muestreo único siempre se toma una muestra de tamaño $n_1 = 15$ para todos los lotes, su curva *ASN* es una

línea en $n_1 = 13 = ASN$. Para el plan de muestreo doble:

$$\begin{aligned} ASN &= n_1 + n_2 \sum_{s=c_1+1}^{r_1-1} b(s, n_1) = 10 + 12 \sum_{s=1}^1 b(s, 10) = \\ &= 10 + 12 \binom{10}{1} p(1-p)^9 = 10 + 120p(1-p)^9 \end{aligned}$$

Gráficamente, obtenemos la Figura 7.10.

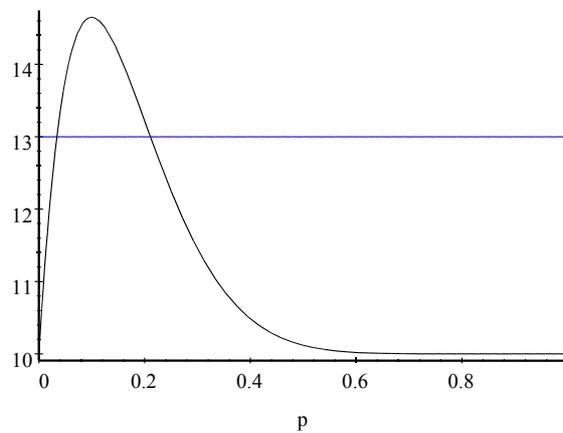


Figura 7.10: Gráfico de comparación del tamaño muestral medio, ASN , para un plan de muestreo único (en azul) y un plan de muestreo doble

La curva muestra que el plan de muestreo doble requiere más inspección para $0.02 \leq p \leq 0.28$. Para lotes de mucha calidad ($p \leq 0.02$) y para lotes de muy baja calidad ($p > 0.28$), el muestreo doble requiere menos inspección total.

Las curvas ASN para planes de muestreo múltiple con más de dos pasos se calculan de la misma manera que en las de muestreo doble.

7.6 El modelo MIL-STD-105E

7.6.1 Desarrollo histórico

Military Standard 105E (MIL-STD-105E) es un sistema de planes de muestreo para la aceptación de atributos que fue inicialmente desarrollado como un conjunto de tablas de inspección para el ejército de Estados Unidos durante la 2ª Guerra Mundial. Esas tablas pasaron por varios niveles de ampliación desde 1942 hasta 1949 hasta que el sistema formal *MIL-STD-105A* se publicó por primera vez en 1950. Posteriormente, se sucedieron diversas modificaciones (*105B-105C-105D*) hasta 1989. En 1989, el sistema se adaptó al *MIL-STD-105E*, que es la versión más actual. Existe una versión civil de este plan militar, el plan *ANSI/ASQC Z1.4 (American National Standards Institute)*, pero supone sólo pequeñas modificaciones de éste. Este sistema también ha sido adoptado por la *International Organization for Standardization* bajo la denominación de *ISO 2859*.

Por su adopción formal por el Gobierno de los Estados Unidos, el muestreo de aceptación fue una herramienta del aseguramiento de la calidad desde finales de los años 50 hasta finales de los 70, cuando la filosofía japonesa de la mejora continua del proceso comenzó a jugar un papel importante en EE.UU. El papel del muestreo para la aceptación fue reexaminado durante los años 80 y como resultado se redujo el énfasis en el muestreo para la aceptación en procesos que están en control estadístico. El cambio en la filosofía del uso del muestreo para la aceptación naturalmente ha reducido el uso del *MIL-STD-105E*, pero este sistema ha sido tan ampliamente utilizado que se debe conocer, además los nuevos sistemas que lo han sustituido (*ANSI/ASQC Z1.4* o *ISO 2859*) son modificaciones sobre éste.

7.6.2 MIL-STD-105E

El *MIL-STD-105E* cubre tres tipos de muestreo: simple, doble y múltiple. Para cada tipo de muestreo existen planes específicos dependiendo del nivel de calidad que el comprador espera del vendedor, esto es, basados en el *AQL*.

Niveles AQL

Los niveles de AQL contenidos en los planes varían desde un 0.01% al 10%, y desde un 10% a 1000%. Los niveles de AQL del 10% o menos se utilizan para contar elementos no conformes; niveles de AQL del 10% y más se utilizan para contar el número de no conformes por 100 unidades.

Hay un patrón repetido en las listas de los AQL . Entre cualesquiera dos AQL que sean sucesivas potencias de 10 hay exactamente cuatro AQL , cada uno seleccionado de forma que represente un múltiplo fijo del AQL inmediatamente inferior a él. Por ejemplo, entre un AQL de 10^{-2} y uno de 10^{-1} , hay AQL de 0.015, 0.025, 0.040 y 0.065 y así sucesivamente. Dado cualquier AQL , digamos p_0 , por este modelo se deduce que hay una constante r para la que los AQL son p_0 , rp_0 , r^2p_0 , r^3p_0 , r^4p_0 y r^5p_0 . Ya que p_0 y r^5p_0 son potencias sucesivas de 10 su razón debe de ser $10 \Rightarrow 10 = \frac{r^5p_0}{p_0} = r^5 \Rightarrow r = 10^{1/5} = 1.5849$. Todos los AQL en el *MIL-STD-105E* se calculan utilizando el factor 1.5849 pero después se redondean a dos dígitos significativos por simplicidad.

Además, los estándares establecen que “la selección o uso de un AQL no implicará que el contratista tenga derecho a suministrar cualquier unidad defectuosa del producto”. En otras palabras, los planes en el *MIL-STD-105E* están diseñados para aceptar lotes en los que $p = AQL$ con una alta probabilidad, pero esto no significa que el suministrador pueda añadir elementos no conformes en un lote mientras mantenga su proporción total de no conformes por debajo del AQL .

Reglas de cambio

Una de las características del *MIL-STD-105E* es que no proporciona simplemente un plan para alcanzar un AQL dado; en vez de esto, proporciona una colección de planes y reglas para cambiar entre estos planes para tener en consideración cambios en la calidad de los lotes suministrados para la inspección. Esto es, para un mismo tamaño de lote y un mismo nivel de calidad aceptable AQL , se especifican tres planes de inspección:

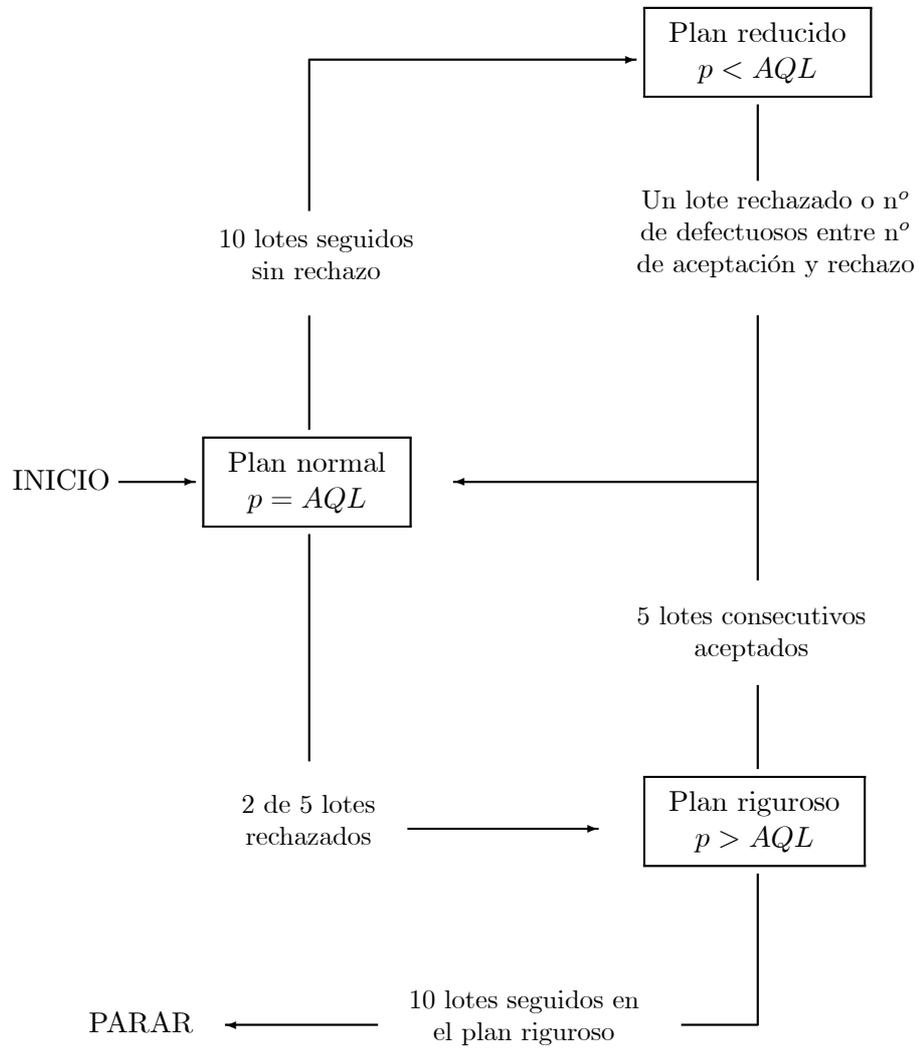
1. Inspección normal: para aquellos casos en los que la calidad que se espera del proveedor es similar al AQL .
2. Inspección reducida: para aquellos casos en los que la calidad esperada es muy alta ($p < AQL$). En este tipo de muestreo, el tamaño muestral es inferior al plan normal.

3. Inspección rigurosa: implica un elevado tamaño muestral. Se utiliza cuando se espera una calidad inferior a la *AQL*.

Cuando se comienza con el *MIL-STD-105E* se elige la inspección normal. Los planes dictados por la inspección normal se utilizan hasta que se rechazan demasiados lotes sucesivos (pasando a la inspección rigurosa) o hasta que se aceptan muchos lotes sucesivos (pasando a la inspección reducida). Por lo tanto, si la calidad de los lotes suministrados comienza a disminuir, la inspección rigurosa hará que los lotes se inspeccionen de forma aún más exhaustiva, proporcionando un incentivo al suministrador para localizar el problema mientras que, simultáneamente, se protege al cliente de lotes de mala calidad. Recíprocamente, si la calidad de los lotes se incrementa, la inspección reducida recompensa al suministrador con una inspección menos fuerte y asegura que el cliente recibe estos lotes de alta calidad. Las reglas que determinan el plan de muestreo se resumen en los siguientes puntos:

- El plan de inspección normal se realizará al comienzo de la tarea de inspección.
- Cambio de plan normal a riguroso: se pasará de inspección normal a rigurosa si dos de cada cinco lotes consecutivos han sido rechazados.
- Cambio de plan riguroso a normal: se pasará de control riguroso a normal cuando se acepten cinco lotes consecutivos.
- Cambio de plan normal a reducido: se pasará de muestreo normal a reducido si no se rechaza ningún lote durante diez lotes seguidos.
- Cambio de plan reducido a normal: se pasará de muestreo reducido a normal si un lote es rechazado. También puede volverse al plan normal cuando el número de defectuosos no lleva ni a aceptar ni a rechazar el lote.
- Si se está en el plan de inspección riguroso durante más de diez lotes, la inspección debe concluir y se debe proponer al vendedor que aumente los niveles de calidad de su producción.

Dichas reglas de cambio se resumen en la Figura 7.11.

Figura 7.11: Reglas de cambio del plan *MIL – STD – 105E*

Dentro de cada uno de esos modos de inspección hay tres *Niveles Generales de Inspección*, denotados por I, II y III, así como cuatro *Niveles Especiales de Inspección*, $s - 1$, $s - 2$, $s - 3$ y $s - 4$. Estos niveles determinan las relaciones entre el tamaño del lote y el tamaño muestral y los elige la autoridad responsable. Salvo que se especifique otra cosa se utiliza el nivel general de inspección II. La diferencia entre los niveles es el grado de discriminación que ofrecen. El nivel I ofrece menos discriminación que el nivel II, el cual ofrece menos discriminación que el nivel III; esto es, las curvas OC de estos planes a estos niveles van decreciendo desde el nivel I al nivel III. Los niveles de inspección especiales se utilizan cuando se necesitan tamaños muestrales más pequeños, ya sea porque la inspección sea destructiva o porque sea muy cara.

Curvas OC

El MIL-STD-105E especifica tamaños muestrales y números de aceptación/rechazo de forma que exista una probabilidad alta de que lotes de una calidad $p = AQL$ o mejores sean aceptados. Sin embargo, para un AQL fijo, P_a no permanece constante si el tamaño del lote cambia. Por ejemplo, para planes con $c \geq 1$, P_a varía de 0.91 a 0.99 si el tamaño del lote crece. Para planes con $c = 0$, $P_a \simeq 0.88$. Esta variación en P_a se diseñó en el sistema con el propósito de incrementar la probabilidad de aceptar los lotes con nivel de calidad AQL cuando el tamaño del lote se incrementa. La razón para hacer esto es que las consecuencias de rechazar lotes son más serias (para el suministrador) si el tamaño del lote se incrementa, luego los esfuerzos se deben hacer para incrementar P_a (para lotes de calidad AQL) en estas circunstancias. Como *MIL-STD-105E* elige tamaños muestrales que se incrementan con el tamaño del lote, P_a (en $p = AQL$) también crece con el tamaño muestral. Esta característica puede ser un poco desconcertante en la práctica porque incluso cambios moderados en los tamaños de los lotes afectan a la probabilidad de aceptar el lote. Para estar seguro del grado de protección que un plan ofrece hay que fijarse en las curvas OC publicadas en el sistema MIL-STD-105E para ese plan en particular.

7.6.3 Implementación

Hay dos pasos fundamentales en la operación del *MIL-STD-105E*:

1. seleccionar el plan apropiado de las tablas publicadas,
2. usar las reglas de cambio cuando la calidad del lote cambia para mejor o para peor.

Desafortunadamente, es fácil olvidar el segundo paso y sólo utilizar el sistema para elegir un plan único que se utiliza después sin mirar los cambios en la calidad del lote. Normalmente esto ocurre cuando se utiliza la inspección normal con nivel II para seleccionar un plan y después, por una razón u otra se sigue utilizando ese plan permanentemente. El problema con esto es que si la calidad del lote comienza a deteriorarse, el cliente acaba aceptando lotes con una calidad más baja de la que se fijó al comienzo.

Para evitar confusión, *MIL-STD-105E* utiliza la notación de “Ac” para “número de aceptación” (*Acceptance number*) y “Re” para “número de rechazo” (*Rejection number*). Para un plan con una muestra, un lote se acepta si el número de elementos no conformes, X , es menor o igual que el “Ac” dado en las tablas y se rechaza si X excede o iguala al número “Re” que dan las tablas. La mayor parte de las veces, “Ac” y “Re” se diferencian sólo en una unidad. En otros casos (inspección reducida, muestreo doble y muestreo múltiple), “Ac” y “Re” pueden diferir en más de una unidad, pero el procedimiento de decisión no cambia: Aceptar si $X \leq Ac$, rechazar si $X \geq Re$.

Los pasos en la implementación del *MIL-STD-105E* se pueden resumir como sigue:

- Paso 1** Determinar el *AQL* (basándose en un acuerdo entre el productor y el cliente).
- Paso 2** Decidir un modo de inspección y un nivel (por defecto para comenzar el sistema: inspección normal, nivel II).
- Paso 3** Determinar el tamaño del lote.
- Paso 4** Utilizar la tabla de Letras de Código y el tamaño del lote para seleccionar la letra código apropiada.
- Paso 5** Decidir el tipo de muestreo: único, doble o múltiple.
- Paso 6** Utilizar la tabla correspondiente al procedimiento de muestreo seleccionado en el paso 5, y el modo de inspección y nivel del paso 2,

para encontrar el/los tamaño(s) muestral(es) y números de aceptación y rechazo para el plan. En los casos donde no exista un plan para un tamaño de lote dado y un *AQL*, hay que tener cuidado para seguir las flechas en las tablas para el plan posible más cercano.

Paso 7 Comenzar utilizando el plan seleccionado en el paso 6 y mantener un registro de lotes aceptados y rechazados de forma que las reglas de cambio puedan aplicarse. Cuando sea necesario un cambio, determinar el modo de inspección y el nivel y repetir los pasos 4-6 para encontrar el plan de muestreo revisado.

Ejemplo 85 *Supongamos lotes de tamaño 400 que son inspeccionados a un AQL del 1.5%. ¿Qué clase de plan es apropiado?. Examinando la tabla de letras de código y el tamaño del lote, encontramos que la letra código para el nivel II de inspección es "H". Buscamos un plan de muestreo único así que procedemos a los planes de muestreo único para inspección normal y encontramos que la letra código H se corresponde a un tamaño muestral de $n = 50$, $Ac = 2$ y $Re = 3$. Por lo tanto, el plan de muestro único consiste en tomar una muestra de tamaño 50 del lote y aceptar el lote si el número de elementos no conformes es menor o igual a 2.*

Ejemplo 86 *En las tablas del MIL-STD-105E algunas combinaciones de los tamaños de los lotes y AQL no tienen un plan de muestreo apropiado y, en estos casos, se deben de seguir las flechas direccionales en las tablas para el plan más cercano que sea posible. Cuando se hace esto hay que tener en cuenta que también cambia el tamaño muestral. Por ejemplo, supongamos que el AQL del ejemplo anterior cambia de 1.5% al 0.065%. Como antes, inspección normal, nivel II, especifica una letra código G, pero examinando el plan de muestreo único para inspección normal sólo encontramos una flecha dirigida al plan de muestreo posible más próximo. Este plan corresponde con la letra código L y especifica que el tamaño muestral debe ser 200 con $Ac = 0$ y $Re = 1$. En otras palabras, dependiendo de las limitaciones impuestas por el tamaño del lote y el AQL, el plan seleccionado puede incluir una gran cantidad de inspección (algunas veces hasta 100%).*

7.7 Planes de muestreo para variables

7.7.1 Introducción

Hasta ahora sólo hemos tratado datos de atributos (específicamente el número de elementos no conformes en una muestra) para muestreo para la aceptación. Ya que muchos procesos generan datos de variables también es deseable tener planes de muestreo para la aceptación basados en medidas de variables. Esta sección presenta los fundamentos de muestreo para la aceptación para variables.

Resaltemos que la mayoría del muestreo para la aceptación se hace utilizando atributos, no variables. Hay varias razones para esto. Primero, las observaciones de variables son casi siempre más costosas de obtener que los datos de atributos, ya que son necesarios instrumentos de medida y operarios para obtener dichas observaciones. El coste también incluye el tiempo necesario para tomar las mediciones de las variables. Estos costos se multiplican rápidamente si se mide más de una característica porque se requiere un plan para cada una de ellas. En cambio, aún en el caso de que se estén examinando varias características para determinar la conformidad o la no conformidad de cada elemento de la muestra, sólo es necesario un plan en el caso de tratar con atributos. Otra característica de los planes de variables es que estos últimos normalmente suponen que la distribución subyacente en los datos es Normal. Esto significa que, antes de establecer un plan para variables, habrá que comprobar esta hipótesis. Finalmente, los planes para variables requieren un cálculo computacional más complicado que el de los planes para atributos.

Como aspecto positivo, señalar que, los planes de muestreo para variables normalmente utilizan tamaños muestrales más pequeños que los planes de atributos para los mismos niveles de protección. Por ejemplo, para lotes de tamaño 500, *MIL-STD-105E* requiere muestras de tamaño 50, y para el mismo nivel de protección, el sistema para variables *MIL-STD-414* requiere muestras de unos 25 elementos.

Además de reducir los tamaños muestrales, trabajar con datos de variables usualmente da una descripción más detallada del proceso que si los datos son de atributos.

7.7.2 Muestreo para la Aceptación con Variables

Planes Forma 1 y Forma 2

Como en el caso de los planes para atributos, los planes de variables operan mediante la comparación de un resumen de los datos de la muestra obtenida con valores de aceptación. Esto es, como en los planes de muestreo de atributos, dado un lote, este se aceptará o se rechazará dependiendo de que el estadístico muestral exceda o no el valor de aceptación. Mientras que en los planes para atributos se cuenta el número de elementos en la muestra que son no conformes, los planes de variables resumen las observaciones obtenidas mediante la media muestral, \bar{X} , de una muestra del lote.

Los datos de variables también se diferencian de los datos de atributos en que se establecen límites de especificación para cada característica del proceso. La posición de la media muestral con respecto a los límites de especificación determinará cuándo un lote tiene una calidad aceptable.

Los estadísticos utilizados en los planes de variables combinan la media muestral, la variabilidad del proceso y los límites de especificación:

$$\frac{LSE - \bar{X}}{\sigma}, \frac{\bar{X} - LIE}{\sigma}, \frac{LSE - \bar{X}}{S}, \frac{\bar{X} - LIE}{S}, \frac{LSE - \bar{X}}{\bar{R}}, \frac{\bar{X} - LIE}{\bar{R}}$$

dependiendo de si se conoce o no la variación del proceso σ . Cuando σ es desconocida se estima, o bien mediante la cuasidesviación típica muestral S , o bien la muestra se divide en submuestras de tamaño 5, y se calcula \bar{R} de esas submuestras. En el último caso no es necesario formar el estimador \bar{R}/d_2 , ya que los valores de aceptación publicados ya tienen en cuenta el factor d_2 .

Por ejemplo, supongamos que un plan de muestreo para variables se utiliza en un proceso para el que se conoce σ y que sólo tiene LSE . Dicho plan requiere que se tome de cada lote una muestra aleatoria de tamaño n . Después, se mide la misma característica en cada uno de los elementos de la muestra, y se calcula \bar{X} para formar el estadístico $\frac{LSE - \bar{X}}{\sigma}$. Por último, se determina un nivel de AQL , se encuentra un valor de aceptación k en una tabla publicada y el lote se acepta siempre que $\frac{LSE - \bar{X}}{\sigma} \geq k$. Es decir, el lote se acepta siempre que la \bar{X} se aleje un número razonable de σ por debajo de LSE .

De forma alternativa, el estadístico $\frac{LSE - \bar{X}}{\sigma}$ se puede utilizar para estimar la proporción \hat{p}_u de productos que caen por encima del límite superior

de especificación. Si comparamos \hat{p}_u con una proporción máxima de no conformes aceptable, M , un lote se acepta siempre que $\hat{p}_u \leq M$. Los valores de M , que dependen del AQL especificado y del tamaño muestral, se encuentran en tablas publicadas.

Para distinguir entre esos dos métodos, los planes de muestreo para la aceptación de variables se clasifican en *forma 1* y *forma 2*. Los planes *forma 1* comparan estadísticos como los anteriores con constantes tabuladas k . Los procedimientos de la *forma 2* convierten los datos muestrales en estimaciones de las proporciones \hat{p}_u y \hat{p}_l de los productos que exceden los límites superior o inferior de especificación, y después comparan estas proporciones con una proporción máxima permitida M que se encuentra en tablas.

Cálculo de valores de aceptación

Para ilustrar cómo se calculan los valores de aceptación, vamos a considerar el tipo más simple de un plan de variables: σ es conocida y se utiliza únicamente el límite superior de especificación. Para obtener el número de aceptación k para un plan de la *forma 1* primero necesitamos especificar un AQL tal y como se ha hecho para los planes para atributos. Supongamos, por ejemplo, que $AQL = 1\%$. Ya que normalmente se quiere que la mayoría de los lotes con $p = AQL$ sean aceptados por el plan, vamos a especificar que esto ocurra con una probabilidad del 95%.

Si hacemos la suposición de que las observaciones del proceso X siguen una distribución Normal, entonces los lotes con $AQL = 0.01$ deben satisfacer la ecuación:

$$P[X > LSE] = AQL = 0.01$$

Tipificamos:

$$P[X > LSE] = P\left[Z > \frac{LSE - \mu}{\sigma}\right] = 0.01$$

donde μ es la media de la característica bajo consideración en lotes con calidad AQL .

Resolvemos esa ecuación

$$\frac{LSE - \mu}{\sigma} = Z_{0.99} = 2.326$$

Como se considera que un lote es aceptable si $\frac{LSE - \bar{X}}{\sigma} \geq k$, el 95% de los lotes con calidad AQL deben satisfacer esta desigualdad, esto es,

$$\begin{aligned} 0.95 &= P \left[\frac{LSE - \bar{X}}{\sigma} \geq k \right] = P \left[\frac{LSE - \mu}{\sigma} + \frac{\mu - \bar{X}}{\sigma} \geq k \right] = \\ &= P \left[\frac{\bar{X} - \mu}{\sigma} \leq \left(\frac{LSE - \mu}{\sigma} - k \right) \right] = P \left[\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \leq \left(\frac{LSE - \mu}{\sigma} - k \right) \sqrt{n} \right] = \\ &= P \left[Z \leq \left(\frac{LSE - \mu}{\sigma} - k \right) \sqrt{n} \right] \end{aligned}$$

De donde

$$\left(\frac{LSE - \mu}{\sigma} - k \right) \sqrt{n} = Z_{0.95} = 1.645$$

Como esta ecuación la deben cumplir los lotes de calidad AQL , es decir los lotes tales que $\frac{LSE - \mu}{\sigma} = 2.326$, podemos sustituir:

$$\frac{2.326 - k}{\sqrt{n}} = 1.645 \Rightarrow k = 2.326 - \frac{1.645}{\sqrt{n}}$$

Es fácil generalizar esa ecuación. Si Z_p denota el valor que en una $N(0, 1)$ deja un área inferior de p , tenemos:

$$k = Z_{1-AQL} - \frac{Z_{P_a}}{\sqrt{n}}$$

donde P_a es la probabilidad especificada de aceptar lotes de calidad AQL .

Nótese que estas ecuaciones son para cualquier tamaño muestral n . Para determinar un único tamaño muestral se debe imponer una condición adicional en la curva característica de operación, curva OC, del plan de variables. Una de estas condiciones es especificar algún nivel de calidad, llamado RQL , que debe ser rechazado por el plan; el RQL se corresponde con el $LTPD$ de los planes de muestreo para atributos. Esto es, se supone que lotes de un nivel de calidad RQL son aceptados con una probabilidad pequeña P'_a específica. Dicha situación se muestra en la Figura 7.12.

Con el mismo razonamiento aplicado anteriormente, se puede demostrar que esta condición adicional resulta en la ecuación:

$$k = Z_{1-RQL} + \frac{Z_{1-P'_a}}{\sqrt{n}}$$

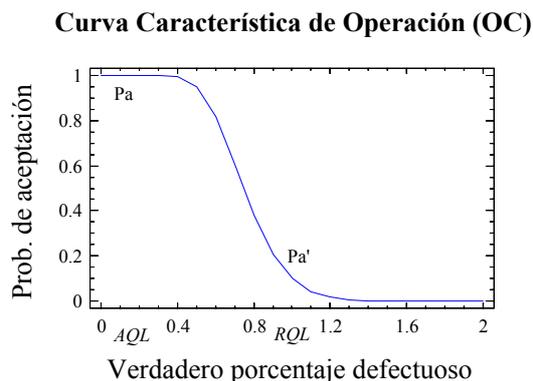


Figura 7.12: Curva OC de un plan de muestreo para variables

Si igualamos esas dos ecuaciones obtenemos:

$$n = \left(\frac{Z_{p_a} + Z_{1-p'_a}}{Z_{1-AQL} - Z_{1-RQL}} \right)^2$$

Normalmente n no sale entero y hay que redondear. Una vez determinado n sustituimos en cualquiera de las ecuaciones anteriores para obtener k . Como n es un redondeo se obtienen dos valores muy próximos de k . Se puede tomar el promedio de esos dos valores, o simplemente seleccionar una de las ecuaciones para obtener un único k .

Por ejemplo, supongamos que añadimos la condición $P'_a = 0.025$ para lotes de calidad $RQL = 0.10$ además del $P_a = 0.95$ para lotes de $AQL = 0.01$.

$$n = \left(\frac{Z_{0.95} + Z_{0.975}}{Z_{0.99} - Z_{0.90}} \right)^2 = \left(\frac{1.645 + 1.960}{2.326 - 1.282} \right)^2 = 11.92$$

Redondeamos a $n = 12$.

Resolvemos las dos ecuaciones de k :

$$k = Z_{0.01} - \frac{Z_{0.05}}{\sqrt{12}} = 2.326 - \frac{1.645}{\sqrt{12}} = 1.851$$

$$k = Z_{0.1} + \frac{Z_{0.025}}{\sqrt{12}} = 1.282 + \frac{1.960}{\sqrt{12}} = 1.848$$

Para el plan de muestreo podemos tomar como k el promedio $\Rightarrow k = 1.8495$ o bien tomar solamente uno de los dos valores.

El mismo tipo de argumento se puede repetir para los casos en los que σ es desconocida y entonces se utilizarían S o \bar{R} para estimar σ . La aproximación es la misma, pero el análisis es algo más complicado porque la variabilidad muestral de S o \bar{R} se debe incluir también en el desarrollo.

7.8 El Modelo MIL-STD-414

7.8.1 Introducción

El sistema *MIL-STD-414*, publicado por primera vez en 1957, se diseñó para muestreo para la aceptación con datos de variables. Como el *MIL-STD-105E*, es un sistema de muestreo para la aceptación en el sentido de que incluye varios planes de muestreo, varios niveles de inspección y un conjunto de reglas para elegir entre los niveles de inspección. El sistema *MIL-STD-414* se desarrolló como un sustituto del *105E* en los casos en los que se deseaban tamaños muestrales pequeños. Esto es especialmente importante cuando, por ejemplo, los elementos muestrales evaluados se destruyen en la inspección. Igual que en el caso del *105E*, el plan *MIL-STD-414* ha sido sustituido por planes civiles como son el *ANSI/ASQC Z1.9* en el caso del American National Standards Institute, o el *ISO 3951* en el caso de la International Standards Organization.

El sistema *MIL-STD-414* es muy parecido al *105E*. Es un sistema basado en el *AQL* con los mismos niveles de inspección (reducida, normal, rigurosa) que el *105E*.

7.8.2 MIL-STD-414

Como la mayoría de los planes de muestreo para variables, el sistema *MIL-STD-414* se basa en la suposición de que las observaciones del proceso siguen una distribución normal. Entonces, antes de utilizar un plan de muestreo de los del sistema, hay que tomar una muestra y ver que realmente se puede suponer que la distribución subyacente es Normal.

El sistema *MIL-STD-414* se puede considerar separado en tres sistemas de muestreo diferentes, cada uno dependiendo del método particular utilizado para estimar la variación del proceso. El sistema incluye conjuntos

separados de planes de muestreo para los casos donde:

1. σ es conocida.
2. σ es desconocida y se estima mediante S .
3. σ es desconocida y se estima mediante \bar{R}/d_2 .

Los procedimientos de la *forma 1* y de la *forma 2* se pueden utilizar con cualquiera de estos tres sistemas, con una excepción: sólo se permiten planes de la *forma 2* para procesos con especificaciones a ambos lados. En la Tabla 7.2 se muestran las posibles combinaciones de la estimación de la variación y los procedimientos de la *forma 1* y de la *forma 2* permitidos en *MIL-STD-414*.

Límites de Especificación	$\hat{\sigma}$	Forma Permitida
A un lado	S \bar{R}/d_2 σ conocida	1 y 2
A dos lados	S \bar{R}/d_2 σ conocida	2

Tabla 7.2: Procedimientos permitidos por el modelo MIL-STD-414

Los planes *forma 2* son los más utilizados porque las reglas de cambio se dan únicamente en términos de porcentajes de no conformes.

Ya que el sistema MIL-STD-414 es un sistema basado en el *AQL*, se debe decidir antes de empezar qué *AQL* se va a utilizar. El sistema MIL-STD-414 contiene sólo un pequeño número de *AQL*, de 0.04 a 15.0. Si estamos interesados en un *AQL* distinto hay que utilizar una tabla de conversión. En la Tabla 7.3 se muestra la tabla de conversión utilizada en el MIL-STD-414.

Por ejemplo, si se desea un *AQL* del 0.75% la Tabla 7.3 muestra que hay que utilizar un *AQL* del 1% para seleccionar uno de los planes del *MIL-STD-414*.

Como en el *105E*, para determinar el tamaño muestral son necesarios el tamaño del lote y el nivel de inspección. En la Tabla 7.4 se muestran las letras de código para diversos tamaños de lote.

Valores específicos de <i>AQL</i>	Utilizar el valor de <i>AQL</i>
- a 0.049	0.040
0.050 a 0.069	0.065
0.070 a 0.109	0.100
0.110 a 0.164	0.150
0.165 a 0.279	0.250
0.280 a 0.439	0.400
0.440 a 0.699	0.650
0.700 a 1.090	1.000
1.100 a 1.640	1.500
1.650 a 2.790	2.500
2.800 a 4.390	4.000
4.400 a 6.990	6.500
7.000 a 10.90	10.00
11.00 a 16.40	15.00

Tabla 7.3: Tabla de conversión

Para una letra de código dada, el tamaño muestral requerido varía dependiendo del nivel de inspección utilizado (es decir normal, reducido o riguroso). Nótese que, como en el *105E* hay cinco niveles de inspección entre los que elegir, pero aquí están numerados I, II, III, IV y V no como en el *105E*. En esencia, los primeros dos niveles (I y II) corresponden a los niveles de inspección especiales de *105E* y los restantes niveles (III, IV y V) a los niveles generales. Es decir, por defecto en, el sistema *MIL-STD-414* se empezará utilizando la inspección normal, nivel IV.

Las tablas de los planes de muestreo del sistema *MIL-STD-414* se utilizan de la misma forma que las del *105E*. Primero, se selecciona la tabla correspondiente al nivel de inspección deseado (normal, reducido o riguroso). Después, la combinación particular de *AQL* y las Letras de Código determinan el tamaño muestral y los valores de aceptación (o porcentajes) para el plan.

7.8.3 Implementación

Como hemos comentado anteriormente, los planes de la *forma 2* son los que se eligen con más frecuencia para trabajar con el *MIL-STD-414*, ya que las reglas de cambio están basadas sólo en términos de porcentajes

Tamaño del lote	Niveles de inspección				
	I	II	III	IV	V
3 a 8	B	B	B	B	C
9 a 15	B	B	B	B	D
16 a 25	B	B	B	C	E
26 a 40	B	B	B	D	F
41 a 65	B	B	C	E	G
66 a 110	B	B	D	F	H
111 a 180	B	C	E	G	I
181 a 300	B	D	F	H	J
301 a 500	C	E	G	I	K
501 a 800	D	F	H	J	L
801 a 1300	E	G	I	K	L
1301 a 3200	F	H	J	L	M
3201 a 8000	G	I	L	M	N
8001 a 22000	H	J	M	N	O
22001 a 110000	I	K	N	O	P
110001 a 550000	I	K	O	P	Q
550001 y más	I	K	P	Q	Q

Tabla 7.4: Letras de código para el MIL-STD-414

de no conformes. Por esta razón, sólo vamos a ver la implementación del *MIL-STD-414*, en la forma 2.

Ejemplo 87 Supongamos que una característica medible de un cierto elemento tiene como límites de especificación 5.000 ± 0.050 cm. y que esos elementos se envían en lotes de tamaño 200. Supongamos además que el límite de especificación inferior es el más crítico de los dos, en el sentido de que elementos que superen el límite superior de especificación pueden ser recuperados acortándolos, mientras que los elementos que caen por debajo del límite inferior de especificación deben ser rechazados. Para reflejar esta diferencia en el plan de muestreo, se seleccionan diferentes AQL en los dos límites de especificación. Supongamos entonces que se selecciona un AQL de 0.80% para el límite de especificación inferior, y un AQL de 2.00% para el límite de especificación superior.

El siguiente paso es trasladar estos AQL a algunos de los permitidos en las tablas del *MIL-STD-414*. De la Tabla 7.3, los AQL superior e inferior

que se utilizarían en el plan son 1.0% y 2.5%, respectivamente.

Como el tamaño del lote es 200, la Tabla 7.4 indica que hay que coger la letra H correspondiente al nivel IV de inspección que se utiliza para comenzar el plan. De la tabla para la inspección normal y rigurosa basada en la variabilidad desconocida: método S, forma 2, el tamaño muestral requerido para el plan es $n = 20$.

La tabla maestra también da el máximo porcentaje de no conformes que corresponde a cada AQL; esto es, $M_l = 2.95\%$ al AQL del 1% y $M_u = 6.17\%$ al AQL del 2.5%. Estos se han de comparar con los porcentajes de no conformes estimados a partir de la muestra. Los estimados a partir de la muestra se hallan primero calculando las cantidades:

$$Q_l = \frac{\bar{X} - LIE}{S}$$

$$Q_u = \frac{LSE - \bar{X}}{S}$$

que se trasladan en porcentajes estimados mediante tablas de conversión especiales basadas en los valores Q_l (o Q_u) y el tamaño muestral n . Por ejemplo, si tenemos la muestra: 5.013, 5.002, 5.032, 4.932, 4.992, 5.042, 5.021, 5.011, 5.015, 5.013, 4.997, 4.966, 4.990, 5.031, 4.968, 5.003, 4.943, 4.952, 4.997, 4.959,

$$Q_l = \frac{\bar{X} - LIE}{S} = \frac{4.994 - 4.95}{0.0310} = 1.42$$

$$Q_u = \frac{LSE - \bar{X}}{S} = \frac{5.050 - 4.994}{0.0310} = 1.81.$$

De la tabla correspondiente, vemos que los porcentajes estimados fuera de los límites de especificación son:

$$\hat{p}_l = 7.49\% \quad \hat{p}_u = 3.05\%$$

Notemos que, de hecho, ninguna de las 20 observaciones de la muestra excede el $LSE = 5.050$ cm. La estimación $\hat{p}_u = 3.05\%$ resulta de la suposición de normalidad e indica que el 3.05% de la población de elementos de la cual se ha sacado la muestra, excede al LSE . La decisión sobre el rechazo o la

aceptación del lote se hace por comparación.

$$\begin{aligned}\hat{p}_l &= 7.49\% > M_l = 2.95\% \\ \hat{p}_u &= 3.05\% < M_u = 6.17\% \\ \hat{p}_l + \hat{p}_u &= 10.54\% > \max[M_l, M_u] = 6.17\%\end{aligned}$$

Dos de estas comparaciones, $\hat{p}_l > M_l$ y $\hat{p}_l + \hat{p}_u > \max[M_l, M_u]$ dan muestras de que el lote se debe rechazar.

En resumen:

Operaciones del sistema de muestreo MIL-STD-414

Paso 1 Determinar el *AQL*. En el caso de límites de especificación dobles se pueden elegir diferentes *AQL*.

Paso 2 Si es necesario, utilizar la tabla de conversión de los *AQL* para obtener *AQL* consistentes con los planes *MIL-STD-414*.

Paso 3 Decidir el modo de inspección y el nivel: al menos que se especifique otra cosa, utilizar la Inspección Normal, nivel IV para comenzar el plan.

Paso 4 Utilizar la tabla de letras de código del tamaño muestral para seleccionar la letra código apropiada.

Paso 5 Decidir si utilizar el plan de la forma 1 o de la forma 2.

Paso 6 Establecer el modelo de estimar la variación del proceso: σ conocida, $\hat{\sigma} = S$ o $\hat{\sigma} = \bar{R}/d_2$.

Paso 7 Utilizar la Tabla Maestra correspondiente a las elecciones anteriores para determinar el tamaño muestral y los valores de aceptación (o porcentajes). En los casos donde un plan no exista para un tamaño de lote dado y un *AQL*, seguir las flechas en la tabla del plan permitido más cercano.

Paso 8 Comenzar el plan seleccionado en el Paso 7 y mantener un registro de lotes aceptados y rechazados para poder aplicar las reglas de cambio.

7.9 Ejercicios

1. Para verificar la calidad de los lotes de 10000 unidades, se realiza el siguiente plan de muestreo para atributos: tomar una muestra de 100 unidades y aceptar el lote si el número de unidades defectuosas es igual o inferior a 5, en caso contrario se rechazará el lote. Si se considera que un lote tiene un nivel de calidad aceptable cuando la proporción de defectuosos en el mismo es 0.02 y que el nivel de calidad para el cual no es aceptable el lote es 0.05, calcular el riesgo del comprador y del vendedor.
2. Un producto es suministrado en lotes de tamaño $N = 3300$. Se especifica un $AQL = 0.10\%$ establece los planes de muestreo único bajo la inspección normal, rigurosa y reducida del sistema de muestreo MIL-STD-105E suponiendo inspección general nivel II.
3. Una empresa ha encargado a cierto suministrador la fabricación de un elevado número de cierto componente electrónico. Para la adquisición del producto, la empresa realiza un muestreo para la aceptación. Los componentes vienen en lotes de tamaño 40000. Se aceptarán aquellos lotes que contengan como máximo un 1% de artículos defectuosos. Para controlar que se verifica dicho nivel de calidad se toma una muestra de tamaño 1500 y se acepta el lote si hay menos de 15 artículos defectuosos (con 15 o más se rechaza, con 14 o menos se acepta). Se pide:
 - (a) Calcular el riesgo del productor y el riesgo del consumidor si el nivel de calidad rechazable es 4 veces el nivel de calidad aceptable.
 - (b) ¿Qué opinión te merece el plan?
4. Un producto es suministrado en lotes de tamaño $N = 10000$. Se especifica un $AQL = 0.1\%$ establece los planes de muestreo único bajo la inspección normal, rigurosa y reducida del sistema de muestreo MIL-STD-105E suponiendo inspección general nivel II.
5. Una empresa compra componentes electrónicos en lotes de 500 unidades. La empresa acuerda con el suministrador un nivel de calidad aceptable de los lotes del 2.5% y un nivel de calidad rechazable del 10%.

El muestreo para la aceptación se hace según el plan MIL-STD-105E con inspección normal y nivel general de inspección II. Se pide:

- (a) Determina el plan de muestreo.
- (b) Determina el riesgo del vendedor.
- (c) Determina el riesgo del comprador.

Apéndice A

Tablas del *Mil-STD-505E*

TABLE I—Sample size code letters

(see 4.9.1 and 4.9.2)

MIL-STD-105E

Lot or batch size	Special inspection levels				General inspection levels		
	S-1	S-2	S-3	S-4	I	II	III
2 to 8	A	A	A	A	A	A	B
9 to 15	A	A	A	A	A	B	C
16 to 25	A	A	B	B	B	C	D
26 to 50	A	B	B	C	C	D	E
51 to 90	B	B	C	C	C	E	F
91 to 150	B	B	C	D	D	F	G
151 to 280	B	C	D	E	E	G	H
281 to 500	B	C	D	E	F	H	J
501 to 1200	C	C	E	F	G	J	K
1201 to 3200	C	D	E	G	H	K	L
3201 to 10000	C	D	F	G	J	L	M
10001 to 35000	C	D	F	H	K	M	N
35001 to 150000	D	E	G	J	L	N	P
150001 to 500000	D	E	G	J	M	P	Q
500001 and over	D	E	H	K	N	Q	R

TABLE II-A—Single sampling plans for normal inspection (Master table)

(see 4.9.3 and 4.9.4)

MIL-STD-105E

Sample size code letter	Sample size	Acceptable Quality Levels (normal inspection)																		
		0.10	0.15	0.25	0.40	0.65	1.0	1.5	2.5	4.0	6.5	10	15	25	40	65	100	150	250	400
A	2	→																		
B	3	→																		
C	5	→																		
D	8	→																		
E	13	→																		
F	20	→																		
G	32	→																		
H	50	→																		
J	80	→																		
K	125	→																		
L	200	→																		
M	315	→																		
N	500	→																		
P	800	→																		
Q	1250	→																		
R	2000	→																		

- Use first sampling plan below arrow. If sample size equals, or exceeds, lot or batch size, do 100 percent inspection.
- ↔ Use first sampling plan above arrow.
- Ac = Acceptance number.
- Re = Rejection number.

TABLE II-B — Single sampling plans for tightened inspection (Master table)

(see 4.9.3 and 4.9.4)

Sample size code letter	Sample size	Acceptable Quality Levels (tightened inspection)																							
		0.010	0.015	0.025	0.040	0.065	0.10	0.15	0.25	0.40	0.65	1.0	1.5	2.5	4.0	6.5	10	15	25	40	65	100	150	250	400
A	2	→																							
B	3	→																							
C	5	→																							
D	8	→																							
E	13	→																							
F	20	→																							
G	28	→																							
H	38	→																							
I	50	→																							
J	65	→																							
K	85	→																							
L	110	→																							
M	150	→																							
N	200	→																							
P	315	→																							
R	500	→																							
S	800	→																							
T	1250	→																							
U	2000	→																							
V	3150	→																							

= Use first sampling plan below arrow. If sample size equals or exceeds lot or batch size, do 100 percent inspection.
 = Use first sampling plan above arrow.
 Ac = Acceptance number.
 Re = Rejection number.

TABLE II-C—Single sampling plans for reduced inspection (Master table)

(see 4.9.3 and 4.9.4)

Sample size ratio levels	Acceptable Quality Levels (reduced inspection)																
	0.010	0.015	0.025	0.040	0.05	0.10	0.15	0.25	0.40	0.5	1.00	1.50	2.50	4.00	6.50	10.00	
A	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
B	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
C	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
D	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
E	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
F	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
G	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
H	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
I	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80
J	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125	125
K	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
L	315	315	315	315	315	315	315	315	315	315	315	315	315	315	315	315	315
M	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
N	800	800	800	800	800	800	800	800	800	800	800	800	800	800	800	800	800

- Use first sampling plan below unless otherwise specified.
- Use first sampling plan below unless otherwise specified.
- Acceptance number.
- Rejection number.
- If the acceptance number has been exceeded, but the rejection number has not been reached, accept the lot. See extensive normal inspection, para 4.1.4.

Bibliografía

- Bertrand, L. H. (1990). *Control de Calidad: Teoría Y Aplicaciones*, Díaz de Santos, Madrid.
- Bissel, D. (1994). *Statistical Methods for SPC and TQM*, Chapman and Hall.
- Burr, I. J. (1967). The effect of nonnormality on constants for \bar{X} and r charts, *Industrial Quality Control*, **23**.
- Crosby, P. B. (1979). *Quality is Free: The Art of Making Quality Certain*, McGraw-Hill, New York.
- Feigenbaum, A. V. (1991). *Total Quality Control*, McGraw-Hill, Inc., New York.
- Montgomery, D. C. (2001). *Introduction to Statistical Quality Control*, John Wiley and Sons, New York.
- Shewart, W. (1931). *Economic Control of Quality of Manufactured Product*, Van Nostrand Reinhold, Princeton (New York).
- Taguchi, G. (1986). *Introduction to Quality Engineering*, Asian Productivity Organization (Distributed by American Supplier Institute Inc., Dearborn, MI).