TEMA 5. MUESTREO Y DISTRIBUCIONES MUESTRALES

APÉNDICE: GRÁFICOS DE CONTROL DE SHEWART

Hay muchas situaciones en las interesa que una variable permanezca constate a lo largo del tiempo. Sin embargo, todos los procesos presentan variación, debida a pequeñas diferencias en las materias primas utilizadas, desajustes de las máquinas, comportamiento del operador, etc. Por tanto, no podemos esperar que una variable permanezca exactamente constante con el paso del tiempo, pero sí que su modelo de variación permanezca estable.

Todo proceso de funcionamiento regular tiene variabilidad debida a dos tipos de causas: asignables y no asignables. Las causas no asignables están presentes siempre, produciendo una variabilidad homogénea y estable que es predecible al ser constante. Las asignables sólo intervienen en determinados momentos, y producen entonces una variabilidad muy grande. Los defectos debidos a causas no asignables aparecen aleatoriamente y la aparición de un defecto no hace más probable la aparición del siguiente. Por el contrario, los defectos debidos a causas asignables se mantienen hasta que eliminemos la causa que los produce.

Causas no asignables:
-Existen muchas, cada una de pequeña importancia.
-Producen una variabilidad estable.
-Es difícil reducir sus efectos.

Ej.: Variaciones debidas a la materia prima, a diferencias de habilidad entre los operarios, a factores ambientales.

Causas asignables:
-Existen un número pequeño pero que produce fuertes efectos.
-Producen una variabilidad imprevisible.
-Sus efectos desaparecen al eliminar la causa.

Ej.: Variabilidad debida a desajuste, errores humanos, lotes defectuosos, fallos de controles.

La responsabilidad de eliminar las causas asignables corresponde al supervisor del proceso: el desajuste de una máquina, el error de un operario, etc., son causas directamente detectables y resolubles dentro del marco del proceso productivo. Sin embargo, la responsabilidad de reducir la variabilidad producida por las causas no asignables, que son la mayoría, corresponde a la Dirección de la Empresa: mejorando la tecnología, cambiando los proveedores y, en general, mejorando el proceso productivo.

Se dice que un proceso está en estado de control cuando la variabilidad que presenta es debida sólo a causas no asignables, es decir, dicha variabilidad es constante a lo largo del tiempo y, por tanto, predecible. En caso contrario se dice que el proceso está fuera de control.

Las gráficas de control (de Shewart) nos permiten distinguir la variación natural presente en un proceso (causas no asignables) de la variación que sugiere que el proceso ha cambiado (causas asignables). Es decir, sirven para controlar un proceso y alertarnos cuando éste ha sido distorsionado y se halla fuera de control. Se utilizan principalmente para controlar el funcionamiento de un proceso industrial, pero también para estudiar la estabilidad de cantidades tan variables como el nivel de ozono en la atmósfera o los índices de audiencia de programas televisivos.

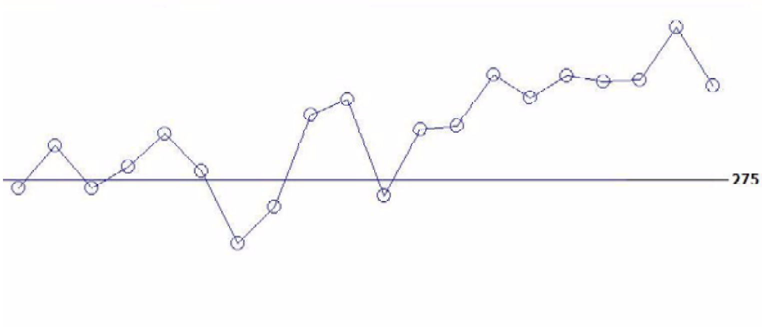
1. Gráficos de Control $\overline{X}$

Ilustraremos este primer tipo de gráfica con un ejemplo.

Ejemplo: El fabricante de un determinado material plástico debe controlar la resistencia a la tensión de sus piezas. Se sabe por experiencia que, cuando el proceso de producción funciona correctamente, la resistencia media de sus piezas es de 275 N y la desviación típica de las lecturas de resistencia es de 43 N. Con el fin de determinar si el proceso de producción se encuentra bajo control, **un operario mide la resistencia a la tensión de una muestra de cuatro piezas cada hora y anota la resistencia promedio**. La siguiente tabla recoge la resistencia a la tensión promedio de cada una de las muestras recogidas a lo largo de 20 horas consecutivas de trabajo. ¿Podemos decir que el proceso se encuentra bajo control? ¿Cómo podemos utilizar estos datos para mantener el proceso bajo control?

Muestra n^o	$\overline{x}$	Muestra n^o	$\overline{x}$	Muestra n^o	$\overline{x}$	Muestra n^o	$\overline{x}$
1	269.5	6	280.4	11	264.7	16	342.6
2	297	7	233.5	12	307.7	17	338.8
3	269.6	8	257.4	13	310	18	340.1
4	283.3	9	317.5	14	343.3	19	374.6
5	304.8	10	327.4	15	328.1	20	336.1

Para estudiar la estabilidad del proceso, representamos las medias de las distintas muestras que se han tomado a lo largo de las 20 horas. Como se ha mencionado anteriormente, el valor esperado para la resistencia a la tensión de las piezas de plástico es de 275 N (blanco para la media del proceso), así que dibujamos una línea horizontal a dicha altura. Las medias de las muestras 12 en adelante son consistentemente mayores que las medias de las primeras muestras, lo que sugiere que la media del proceso va aumentando con el tiempo.



Pero quizás este cambio refleje la variación natural en el proceso. Para averiguarlo, haremos algunas consideraciones teóricas.

La variable aleatoria en consideración es:

$$X = \text{"Tensión de la malla metálica de un monitor seleccionado al azar"},$$

cuyan distribución no se alejará mucho del modelo Normal.

De esta variable aleatoria sabemos que, cuando el proceso de fabricación funciona adecuadamente, la media es $\mu = 275$ y la desviación típica es $\sigma = 43$.

Por otra parte, sabemos que la media muestral para muestras de tamaño 4,

$$\overline{X} = \frac{X_1 + X_2 + X_3 + X_4}{4}$$

sigue una distribución aproximadamente Normal (Teorema Central del Límite) con media $\mu = 275$ y desviación típica $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{43}{\sqrt{4}} = 21.5$

En general, la probabilidad de que la media muestral esté en el intervalo $(\mu - 3\sigma_{\overline{X}}, \mu + 3\sigma_{\overline{X}})$ es aproximadamente 0.997:

$$P(\mu - 3\sigma_{\overline{X}} \leq \overline{X} \leq \mu + 3\sigma_{\overline{X}}) = 0.997$$

Por tanto, si alguna media aritmética $\overline{x}$ no cae dentro del intervalo $(\mu - 3\sigma_{\overline{X}}, \mu + 3\sigma_{\overline{X}})$, es evidencia de que el proceso está fuera de control, es decir, que la distribución del proceso ha cambiado (la media μ ha cambiado).

En nuestro ejemplo, tenemos que:

$$P(275 - 3 \cdot 21.5 \leq \overline{X} \leq 275 + 3 \cdot 21.5) = P(210.5 \leq \overline{X} \leq 339.5) = 0.997$$

de manera que si alguna de las medias muestrales $(\overline{x})$ de tamaño 4, no cae dentro del intervalo $(210.5, 339.5)$, es evidencia de que el proceso está fuera de control (la media μ ya no es 275).

Nota: El intervalo que determina si el proceso está o no bajo control da lugar a los denominados **límites de control.**

Construcción del gráfico $\overline{X}$:

1. Se traza una línea horizontal a la altura del "blanco" (valor que tendría la media del proceso si estuviera bajo control)

2. Se trazan dos líneas horizontales a la altura de los límites de control, que se calculan de la siguiente forma:

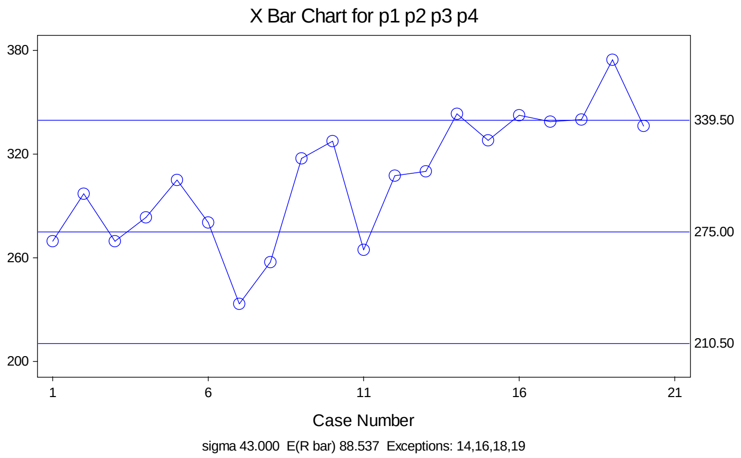
$$\begin{aligned} \text{Límite Superior de Control} &= \mu + 3\sigma_{\overline{X}} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} \\ \text{Límite Inferior de Control} &= \mu - 3\sigma_{\overline{X}} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} \end{aligned}$$

3. Se representan en el gráfico anterior las medias aritméticas de las muestras seleccionadas. Si alguna cae fuera de los límites de control, entonces el proceso ha cambiado y debemos averiguar las causas.

Para el ejemplo anterior, las líneas del gráfico de control vienen dadas por:

$$\begin{aligned} &\text{Límite Superior de Control} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} = 275 + 3 \cdot \frac{43}{\sqrt{4}} = 339.\,5 \\ &\text{Línea Central} \qquad\qquad\quad = \mu = 275 \\ &\text{Límite Inferior de Control} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} = 275 - 3 \cdot \frac{43}{\sqrt{4}} = 210.\,5 \end{aligned}$$

y el gráfico de control $\overline{X}$ sería el siguiente:



Además de la presencia de puntos más allá de los límites de control existen otros indicativos de estado "fuera de control" cuando los puntos exiben algún patrón no aleatorio de comportamiento.

En ocasiones resulta útil representar un par de líneas internas a las de control para utilizarlas como **líneas de alerta**, se suelen situar a ± 2 desviaciones típicas de la media, por lo que entre ambas se encontrará aproximadamente el 95.44% de las observaciones. Estos límites de alerta se utilizan para comenzar a tomar medidas preventivas.

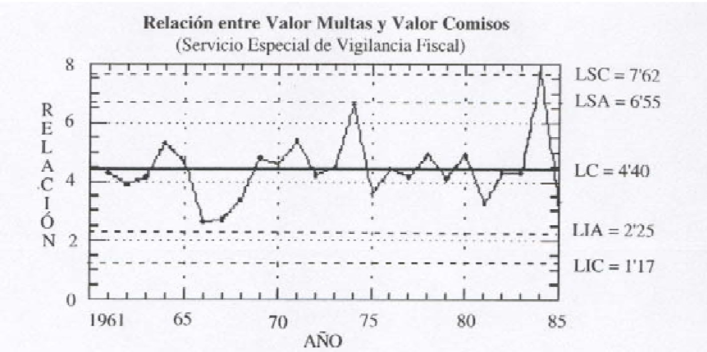


Fig. 19.2.1: Ejemplo de Gráfico de Control.

En general, 8 o más puntos consecutivos por encima o debajo de la media, lo que se denomina **racha**, o en orden creciente o decreciente, denominados **tendencia**, se consideran indicativos de anormalidad, ya que la probabilidad de que aparezcan al azar cada una de esas configuraciones es de $(1/2)^8$.

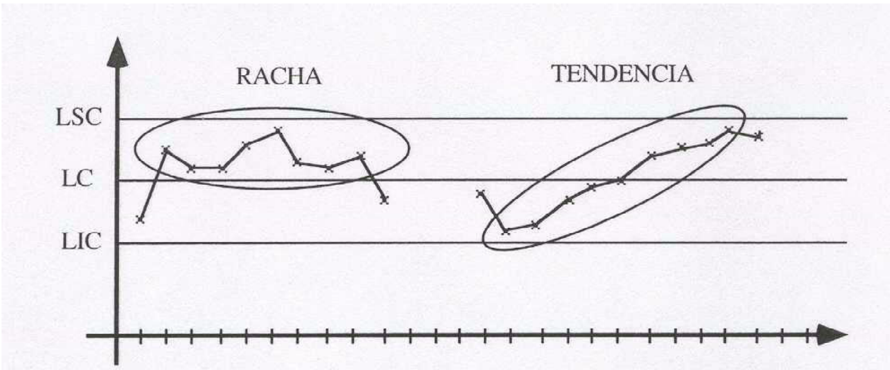


Fig. 19.2.2: Alteraciones que indican anormalidad.

2. Gráficas de Control p

Las gráficas de control p se suelen utilizar para controlar la proporción "p" de artículos que no se ajustan a un conjunto de especificaciones. El procedimiento a seguir es totalmente análogo al de las gráficas de control $\overline{X}$, teniendo en cuenta que el estadístico proporción muestral $\widehat{P} = \dfrac{X_1 + ... + X_n}{n}$ puede aproximarse por una distribución $Normal\left(\mu = p, \sigma = \sqrt{\frac{p(1-p)}{n}}\right)$ (para n suficientemente grande).

Ejemplo: En una fábrica de rodamientos, se comprueba diariamente una muestra de 400 unidades, determinando si se trata de un artículo defectuoso o por el contrario verifica las especificaciones requeridas. En la siguiente tabla aparecen las proporciones observadas de artículos defectuosos durante 16 días.

Día n^o	1	2	3	4	5	6	7	8
$\widehat{p}$	0.115	0.16	0.13	0.1225	0.1	0.1225	0.19	0.115
Día n^o	9	10	11	12	13	14	15	16
$\widehat{p}$	0.1	0.16	0.1675	0.1225	0.1375	0.1975	0.1525	0.1675

Si los fabricantes consideran como proporción estable de defectuosos $p = 0.1$, ¿cómo podemos utilizar estos datos para mantener el proceso bajo control?

Igual que hicimos en el caso anterior, trazamos una recta horizontal a la altura de nuestro "blanco" $p = 0.1$ y calculamos los límites de control:

$$\begin{aligned}\text{Límite Superior de Control} &= p + 3\sigma_{\widehat{P}} = p + 3 \cdot \sqrt{\frac{p(1-p)}{n}} \\ \text{Límite Inferior de Control} &= p - 3\sigma_{\widehat{P}} = p - 3 \cdot \sqrt{\frac{p(1-p)}{n}}\end{aligned}$$

Después representamos las proporciones observadas en cada uno de los 16 días. Cada vez que estos puntos se salen del intervalo formado por los límites de control, significa que el proceso de fabricación de placas base ha sufrido una perturbación y ha cambiado su distribución. Por tanto, ese mismo día debemos ajustar la máquina.

Construcción del gráfico p :

1. Se traza una línea horizontal a la altura del "blanco" (valor que tendría la proporción del proceso si estuviera bajo control)

2. Se trazan dos líneas horizontales a la altura de los límites de control, que se calculan de la siguiente forma:

$$\begin{aligned}\text{Límite Superior de Control} &= p + 3\sigma_{\widehat{P}} = p + 3 \cdot \sqrt{\frac{p(1-p)}{n}} \\ \text{Límite Inferior de Control} &= p - 3\sigma_{\widehat{P}} = p - 3 \cdot \sqrt{\frac{p(1-p)}{n}}\end{aligned}$$

3. Se representan en el gráfico anterior las proporciones observadas. Si alguna cae fuera de los límites de control, entonces el proceso ha cambiado y debemos averiguar las causas.

Para el ejemplo anterior, las líneas del gráfico de control vienen dadas por:

$$\text{Límite Superior de Control} = p + 3 \cdot \sqrt{\tfrac{p(1-p)}{n}} = 0.1 + 3 \cdot \sqrt{\tfrac{0.1(1-0.1)}{400}} = 0.145$$

$$\text{Línea Central} \qquad\qquad = p = 0.1$$

$$\text{Límite Inferior de Control} = p + 3 \cdot \sqrt{\tfrac{p(1-p)}{n}} = 0.1 - 3 \cdot \sqrt{\tfrac{0.1(1-0.1)}{400}} = 0.055$$

y el gráfico de control p sería el siguiente: