

## 1. CURVAS DE OPERACIÓN DEL GRÁFICO DE CONTROL

Las curvas de operación son útiles para evaluar a priori la sensibilidad del gráfico de control frente a desplazamiento del proceso o al cambio de su variabilidad. De estas curvas se deduce que el hecho de que “los puntos estén dentro de los límites” proporciona muy poca seguridad de que el proceso se encuentre realmente bajo control estadístico y pone de manifiesto las limitaciones de esta técnica. La aplicación de gráficos de control debe continuarse con técnicas estadísticas avanzadas encaminadas al conocimiento de los parámetros que influyen en el proceso (CTP en terminología 6 Sigma) y a la reducción de la variabilidad (o lo que es lo mismo, al aumento de los índices de capacidad).

### 1.1 CURVAS DE OPERACIÓN DEL GRÁFICO DE MEDIAS

Si se mide el desplazamiento del proceso en unidades de la desviación típica y la deformación mediante la relación existente entre la desviación típica final y la inicial, es decir:

$$d = \frac{\mu_1 - \mu_0}{\sigma_0} \quad \lambda = \frac{\sigma_1}{\sigma_0}$$

La probabilidad de que la media muestral caiga dentro de los límites de control en función de  $(d, \lambda)$  es:

$$\begin{aligned} P(d, \lambda) &= P\left(\frac{\mu_1 - LCI}{\sigma_1/\sqrt{n}} \leq \frac{\bar{x} - \mu_1}{\sigma_1/\sqrt{n}} \leq \frac{LCS - \mu_1}{\sigma_1/\sqrt{n}}\right) \\ &= P\left(\frac{(d+3)}{\lambda} \leq \frac{\bar{x} - \mu_0 - d\sigma_0}{\sigma_1/\sqrt{n}} \leq \frac{(3-d)\sqrt{n}}{\lambda}\right) \end{aligned}$$

Si analizamos ahora los dos efectos por separado y dibujamos las curvas para distintos tamaños de muestras, vemos que el gráfico de medias es sensible al desplazamiento y muy poco sensible al cambio de la variabilidad (ver Figura 12 y Figura 13).

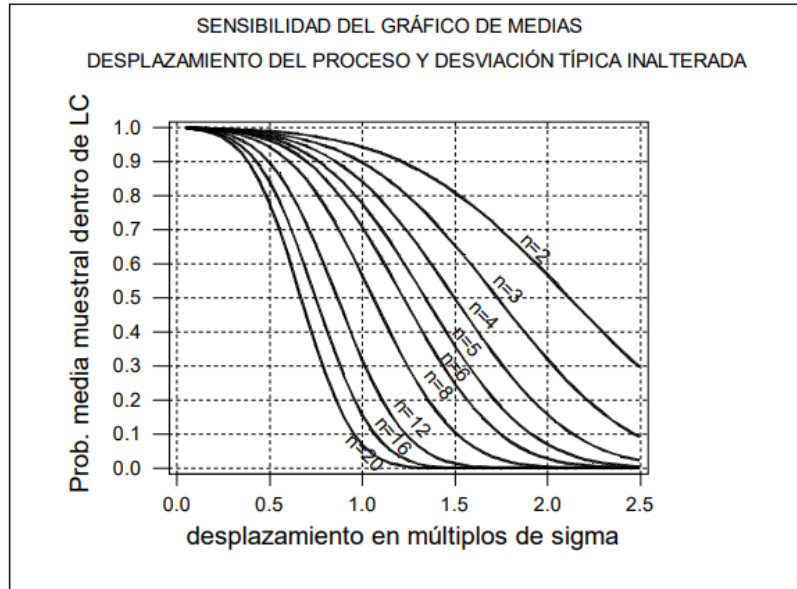


Figura 12: Curva de operación del Gráfico R al desplazamiento de la media

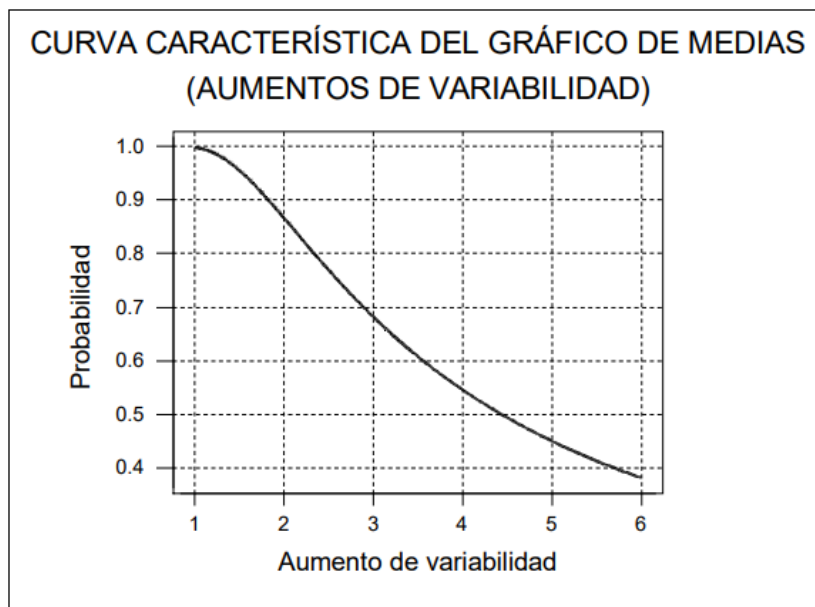


Figura 13: Curva de operación del Gráfico X al aumento de la variabilidad

## 1.2 CURVA DE OPERACIÓN DEL GRÁFICO DE DISPERSIÓN

La construcción de la curva de operación del gráfico de recorridos se hace aprovechando el estadístico  $w = R/\sigma$ . La ecuación de la curva será:

$$P(\lambda)P(D_1\sigma_0 \leq w\lambda\sigma_0 \leq D_2\sigma_0)$$

Para el caso del gráfico S, la probabilidad de que una muestra caiga dentro de los límites de control puede hacerse utilizando el estadístico:

$$\frac{ns^2}{\sigma^2} = X_{n-1}^2$$

Entonces la ecuación de la curva de operación es:

$$P(\lambda) = P\left(\frac{B_1^2\sigma_0^2}{\sigma_1^2}n \leq \frac{s^2}{\sigma_1^2}n \leq \frac{B_2^2\sigma_0^2}{\sigma_1^2}n\right) = P\left(\frac{B_1^2}{\lambda^2}n \leq \frac{s^2}{\lambda^2\sigma_1^2}n \leq \frac{B_2^2}{\lambda^2}n\right)$$

Se puede comprobar que el comportamiento es similar para muestras pequeñas. Sin embargo, para muestras  $n > 6$  es más eficaz el gráfico S.

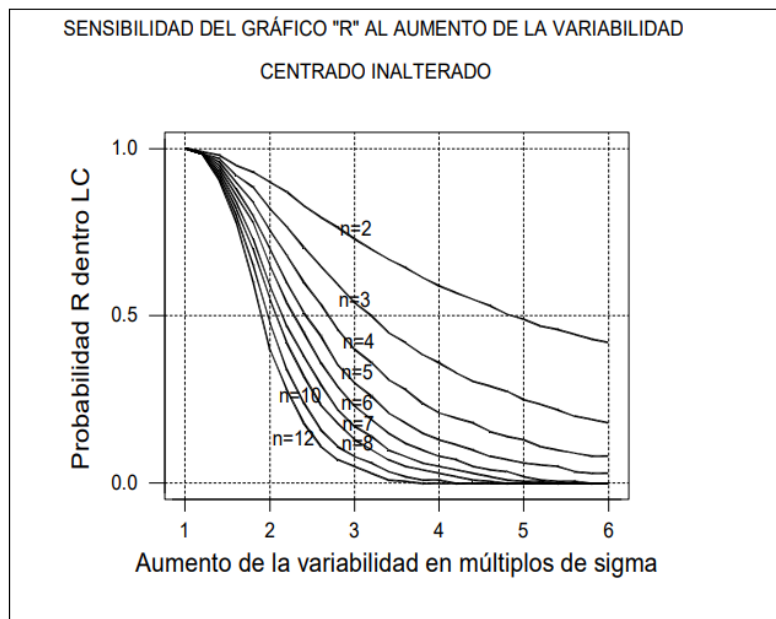


Figura 14: Curva de operación del Gráfico R al aumento de la variabilidad

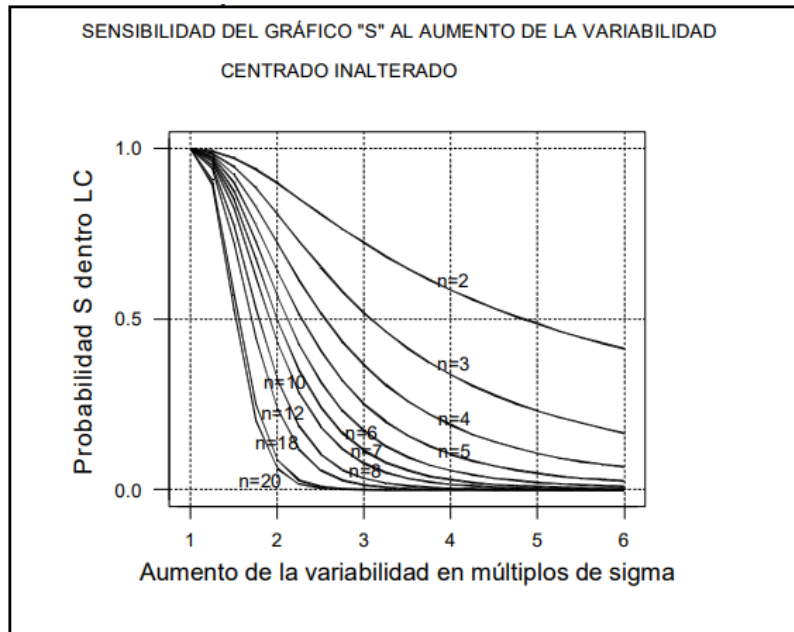


Figura 15: Curva de operación del Gráfico S al aumento de la variabilidad

## 2. INTERPRETACIÓN DE LAS INDICACIONES DEL GRÁFICO DE CONTROL

La indicación más clara de que el proceso está fuera de control es que alguno de los puntos esté fuera de los límites. Además, el aspecto de los gráficos puede indicar anomalías en el proceso. En efecto, si dividimos el gráfico en zonas A, B y C, ver Figura 19, el porcentaje de puntos contenidos en cada una de ellas deberá acercarse sensiblemente al área relativa de la campana de Gauss cubierta por cada una de ellas.

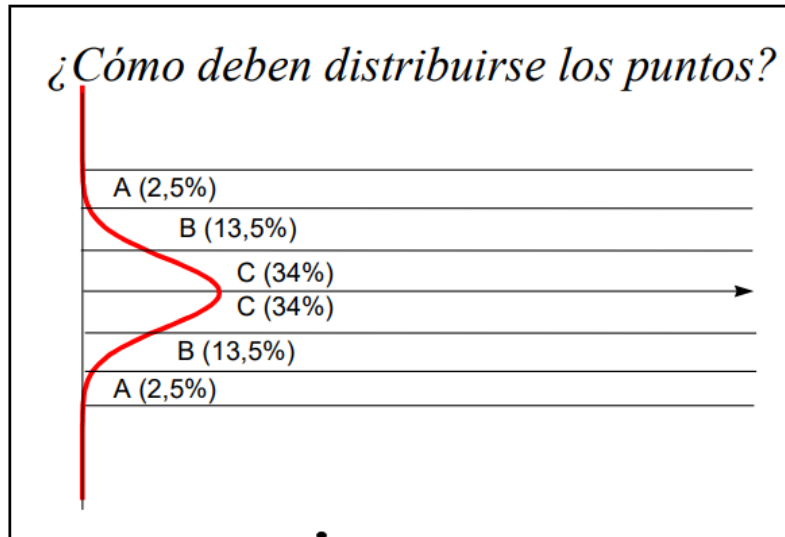


Figura 16: Distribución de los puntos procedentes de una distribución normal

En la Tabla 6 se incluyen algunas reglas utilizadas frecuentemente para analizar el gráfico de medias y establecer si el proceso se encuentra fuera de control. En la columna de la derecha se indica la probabilidad de que se diera esta circunstancia en el caso de un proceso normal, centrado (lo que daría lugar a una decisión errónea). En general, un procedimiento de SPC debe explicitar qué indicaciones deben considerarse como síntoma de proceso fuera de control.

Regla	Descripción	Probabilidad
1	Un punto está fuera de los límites de control	0,27%
2'	De tres puntos consecutivos, dos o más están en una de las zonas A (la misma).	0,15%
2	Dos puntos consecutivos están en una de las zonas A (la misma).	0,05%
3	Racha de cinco puntos en la que cuatro están en la misma zona A o B.	0,27%
3'	Racha de cuatro puntos consecutivos que están en la misma zona A o B.	0,06%
4	Racha de siete puntos consecutivos por encima de la línea central.	0,78%
4'	Racha de seis puntos consecutivos por encima de la línea central.	1,56%
4''	Racha de ocho puntos consecutivos por encima de la línea central.	0,39%
5	Racha de seis puntos consecutivos se presentan de manera monótona creciente.	0,14%
6	De diez puntos consecutivos, existe una racha de 8 que se presentan de manera monótona creciente.	0,07%
6'	De nueve puntos consecutivos, existe una racha de 8 que se presentan de manera monótona creciente.	0,02%
6''	De nueve puntos consecutivos, existe una racha de 7 que se presentan de manera monótona creciente.	0,42%
7	Dados dos puntos consecutivos, el segundo está al menos cuatro desviaciones típicas por encima del primero (por ejemplo, el primero está en la zona A inferior y el segundo en la B o A superior).	0,23%
8	Racha de quince puntos en la zona C	0,30%
9	Racha de catorce puntos alternativos ascendente descendente	?

Tabla 6

A continuación, veamos cómo se pueden interpretar estas indicaciones:

- a) **Reglas 1, 2, 3, 3', 4, 4', 4'' y 7.** Si el gráfico de medias presenta alguna de estas indicaciones y el gráfico S o R no presenta anomalías, lo más probable es que el proceso se encuentre desplazado. Si el gráfico S o R presenta también indicaciones,

entonces es proceso ha aumentado la variabilidad y posiblemente también se ha desplazado.

- b) **Reglas 5, 6, 6', 6''.** Lo más probable es que el proceso esté desplazándose.
- c) **Regla 7.** Si el gráfico S o R no presenta anomalías, lo más probable es que el proceso sea muy inestable en cuanto a su centrado. Si el gráfico S o R presenta también indicaciones, entonces es probable que el proceso haya aumentado la variabilidad.
- d) **Regla 8.** que Puede ser síntoma de que se ha reducido la variabilidad del proceso si el gráfico S o R también muestra una reducción de la variabilidad muestral. Si el gráfico S o R no muestra una reducción de la variabilidad muestral, puede ocurrir que se estén mezclando dos poblaciones distintas, siendo la distribución resultante bimodal (dos jorobas).
- e) **Regla 9.** Es síntoma de inestabilidad del proceso.

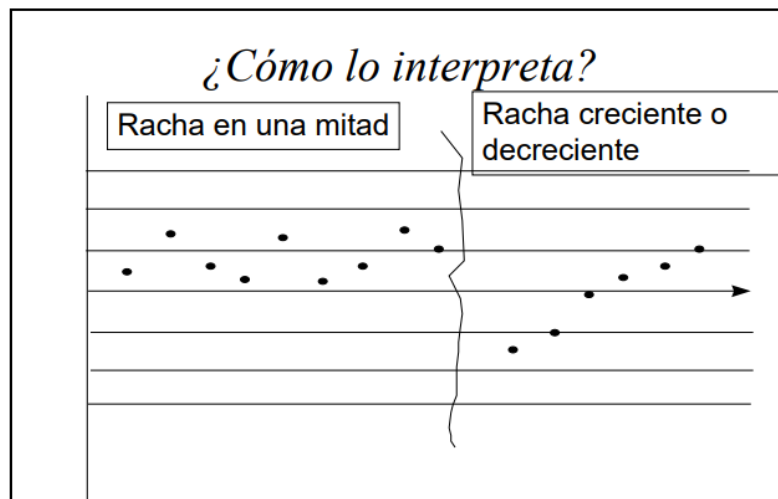


Figura 17: Tendencias que indican causas especiales

Para el caso del gráfico S o R se puede establecer un conjunto de reglas parecidas, pero es necesario tener en cuenta lo siguiente:

- a) La interpretación de la indicación ha de hacerse en función a la naturaleza del gráfico S o R, es decir, no indican derivas en la media del proceso sino aumento o disminución en la variabilidad de este.
- b) La distribución de S o R no es normal (aunque lo sea el proceso) ni tampoco es simétrica, por lo que las probabilidades indicadas en la columna de la derecha no se aplican a este caso.

### **3. ESTUDIO DE CAPACIDAD DE PROCESOS**

#### **10.1 ¿QUÉ ES UN ESTUDIO DE CAPACIDAD DE PROCESO?**

Un estudio de capacidad de proceso tiene por objeto de conocer:

- Distribución estadística que lo describe (normal o no normal)
- El patrón de variabilidad del proceso y principales factores relacionados con la variabilidad.
- Comprender los fenómenos físicos y tecnológicos importantes para el proceso

Se realiza tomando muestras de la producción. A partir de aquí existen muchas diferencias en cuanto al modo de tomar las muestras. En particular es frecuente que se requiera que el proceso esté en control estadístico. Este requisito es lógico cuando se pretende calcular los índices  $C_p$  y  $C_{pk}$  de variabilidad a corto plazo. En general, cuando se desee estimar la variabilidad total del proceso (conocida también como variabilidad a largo plazo), las muestras deben ser representativas de la producción, por ejemplo, debe incluir producciones con lotes distintos de materia prima, realizada por operarios diferentes, en varios turnos, etc. Obviamente, según sea el modo en el que se tomen las muestras los resultados serán muy diferentes. En los párrafos siguientes se dan algunas directrices que pueden ayudar a diseñar un estudio de capacidad de proceso y a interpretar sus resultados.

#### **10.2 ¿CÓMO SE REALIZA UN ESTUDIO DE CAPACIDAD DE PROCESO?**

En la realización de un estudio de capacidad de proceso es importante realizar los siguientes pasos.



### **10.2.1 Comprender los fenómenos físicos y tecnológicos importantes para el proceso**

Naturalmente de un proceso no se conoce “todo”, ya que si fuera así no tendría sentido realizar otro estudio sobre él. Sin embargo, normalmente se saben algunos aspectos que deben tenerse en cuenta, ya que influirán en los datos. Por ejemplo, si se trata de un proceso de mecanizado, es esperable que el desgaste de la herramienta produzca un desplazamiento de la media, etc. Como resultado de la revisión de los conocimientos disponibles se puede llegar a lo siguiente:

- Establecimiento de las posibles fuentes de variabilidad del proceso (por ejemplo, una deriva temporal en el caso del mecanizado, variabilidad entre lotes, variabilidad posicional en la pieza, etc.).
- Verosimilitud de esperar una distribución normal de los datos o si por el contrario existe alguna razón (por ejemplo, asimetrías) que sugieren otro tipo de distribución.

### **10.2.2. Definición de la estratificación de las medidas a tomar**

En función de las consideraciones anteriores, se debe definir el Grupo Homogéneo Racional (GHR) del que se tomarán las muestras y que se espera que estén libres de causas especiales, de manera que la variación dentro de estos grupos pueda ser representativa de la variación del proceso cuando se eliminen todas las causas especiales. Por ejemplo, si se han identificado como posibles fuentes de variabilidad de un proceso de mecanizado de un vástago cilíndrico:

- El desgaste de la herramienta (asociado al tiempo de trabajo).
- El proveedor de la materia prima.
- El lote suministrado por el proveedor.
- Si el diámetro medido corresponde a la primera cara o a la segunda (posición en la pieza).

En este caso cuando se mida la muestra es necesario identificar la hora en la que se fabricó la pieza, qué proveedor y qué lote de materia prima se empleó y a qué cara corresponde la medida, ya que de otro modo será imposible cuantificar las fuentes de variabilidad. En todos los casos debe pensarse que la propia medida puede ser una fuente

de variabilidad más, por lo que deben tomarse las siguientes precauciones:

- Debe definirse el procedimiento de medida con todo detalle. En un caso simple como este, podría bastar con indicar a un técnico de taller que la medida debe tomarse con un micrómetro calibrado y medirse dos diámetros a 90° y calcular la media. Si la medida fuese más compleja, sería necesario establecer un procedimiento escrito.
- Si no se conoce la repetitividad y reproducibilidad del método de medida a emplear, debe realizarse un estudio R&R (repetitividad y reproducibilidad). Los estudios R&R también se conocen como estudios MSA (Measurement System Analysis)

Como resultado de esta fase debe prepararse al menos una plantilla de recogida de datos que evite que se produzcan errores en la recogida o registro de estos.

### **10.2.3. Estabilizar el proceso**

Para tomar los datos, el proceso debe operar de manera que el GRH se encuentre libre de otras posibles causas especiales. Para que tenga sentido la aplicación de los gráficos de control, el proceso ha de tener una estabilidad suficiente que, aun siendo aleatorio, permita un cierto grado de predicción. En general, un proceso caótico no es previsible y no puede ser controlado. A estos procesos no se les puede aplicar el gráfico de control ni tiene sentido hablar de un estudio de capacidad. Un proceso de este tipo debe ser estudiado mediante herramientas estadísticas avanzadas hasta que el grado de conocimiento empírico obtenido sobre el mismo permita estabilizarlo.

### **10.2.4. Toma de los datos**

Generalmente la toma de los datos se extiende en el tiempo y se toman a través de un operario. Este intervalo de tiempo puede ser largo si se pretende capturar fuentes de variabilidad cíclica. Consecuentemente las posibilidades de errores por “malos entendidos” son altas. Por esta razón es muy importante definir una plantilla de toma de datos lo más explícita posible y no esperar a que estén todos los datos tomados para “echar un vistazo”. Debe

evitarse la situación de descubrir que los datos que se han ido tomando durante semanas...no valen para nada.

#### **10.2.5. Identificación del patrón de variabilidad**

Uno de los objetivos del estudio establecidos en el párrafo 10.1 es identificar el patrón de la variabilidad y los principales factores relacionados con la variabilidad. Para ello resultan de la máxima utilidad herramientas como el gráfico Multivari y la técnica de análisis de componentes de la varianza.

#### **10.2.6. Comprobación de la normalidad de los datos**

Una vez obtenidos los datos se debe comprobar la normalidad. Al menos deben ser normales los datos correspondientes al mismo GHR (en caso de que los datos en su conjunto no sean normales, pero si lo sean los datos de los GHR, quiere decir que alguno de los criterios seguidos en la estratificación de las muestras es una causa especial). En caso de que por consideraciones tecnológicas o el análisis de los datos indique que no son normales, debe realizarse una transformación normalizadora.

Es preciso tener en cuenta que, si los datos no son normales, no son válidas las predicciones de fracción defectuosa realizadas en el estudio. En el párrafo 11 se indica cómo proceder en casos en los que los datos no sean normales.

Para comprobar la normalidad se pueden emplear los contrastes de ajuste habituales ( $\chi^2$ , Kolmogorov-Smirnov, etc.). La prueba de normalidad más sencilla y utilizada en ingeniería de calidad es la que se basa en el "papel probabilístico normal". Este papel tiene la escala de ordenadas modificada de manera que las probabilidades acumuladas de datos de una muestra procedente de una distribución normal resulten alineadas.

#### **APLICACIÓN MINITAB**

Dentro del menú **Graph->Probability Plot**, Minitab cuenta con contrastes gráficos de ajuste para varias distribuciones. Si se selecciona la distribución normal y se representan los datos del Ejemplo 3 (ver Figura 21) se observa que alguno de los valores de los extremos está fuera de las bandas de confianza del 95%, lo que es un indicio de falta de normalidad.

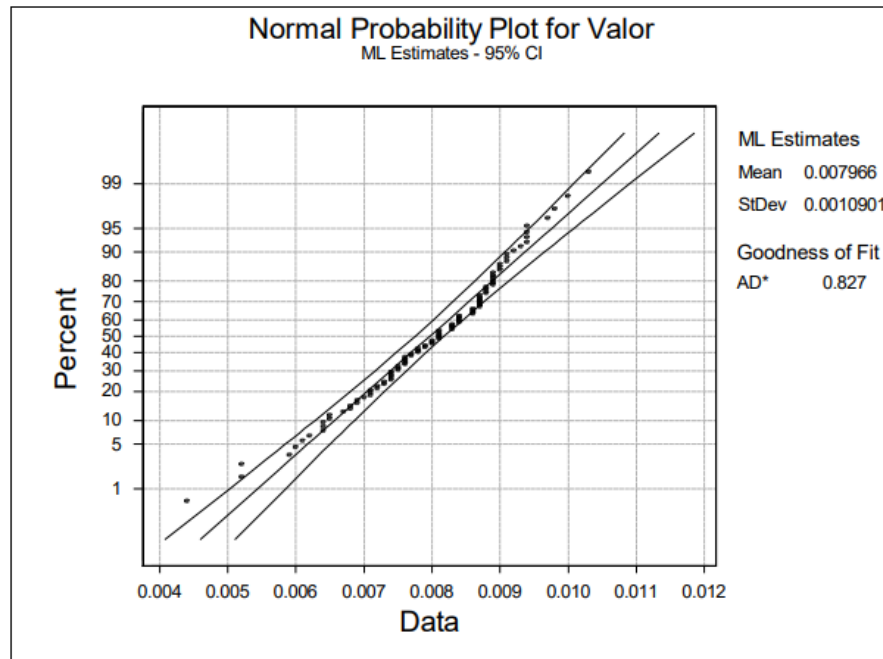


Figura 18: Graph->Probability- Pilot->Normal

### 10.2.7. Análisis e interpretación de los datos

A continuación, debe comprobarse la estabilidad del proceso para lo que debe realizarse un gráfico de control a las medias de los Grupos Homogéneos Racionales comprobando que no existen puntos fuera de control. En caso de que exista algún punto fuera de control, debe estudiarse la causa para su eliminación.

Una vez hecho esto, debe realizarse la estimación de la desviación típica. En general deben seguirse los siguientes criterios:

- Si el tamaño de los grupos es igual o mayor que 10, no se empleará nunca el método del recorrido medio.
- Si se sospecha que la desviación típica es sensiblemente constante, se puede emplear el método de la desviación típica promediada ("pooled standard deviation").

## APLICACIÓN MINITAB

Dentro del menú **Stat->Control Charts-> Xbar-R**, Si se selecciona la columna en la que están los datos y se indican el tamaño del lote = 5, se obtiene la Figura 19. En ella puede apreciarse dos puntos fuera de control, lo que indica que la variabilidad a corto plazo, representada por los índices  $C_p$  y  $C_{pk}$  no debe utilizarse para cuantificar la variabilidad total del proceso.

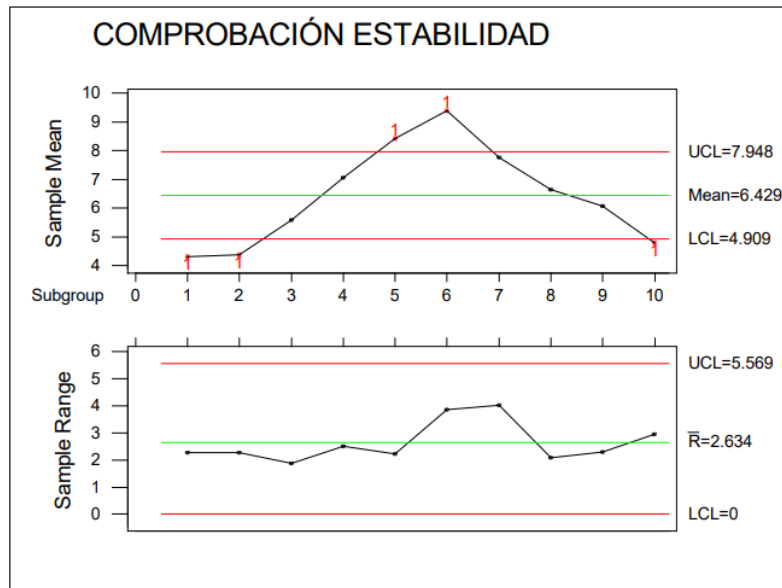


Figura 19: Stat-> Control Charts -> Xbar-R

### Ejemplo 4: Métodos de estimación de la variabilidad. Cálculo de los índices

La variabilidad inherente del proceso ( $\sigma_{CP}$ ) puede estimarse a través del recorrido muestral medio  $\bar{R}$ , (para tamaños de muestra  $n < 10$ ), y también puede hacerse a través de la media ponderada de las desviaciones típicas muestrales. Veamos un ejemplo con los datos de la Tabla 7).

Tolerancia superior: 10.00  
 Tolerancia superior: 5.00

	Muestra 1	Muestra 2	Muestra 3	Muestra 4	Muestra 5	Muestra 6	Muestra 7	Muestra 8	Muestra 9	Muestra 10
Pieza 1	5.35	4.55	5.05	7.41	9.23	8.92	9.08	7.18	5.35	4.45
Pieza 2	3.96	5.85	5.41	8.02	6.99	11.53	6.95	6.75	5.98	4.57
Pieza 3	4.7	4.13	6.18	7.19	9.17	9.37	7.13	6.66	5.03	4.88
Pieza 4	4.4	3.59	6.54	7.11	8.58	9.37	9.81	5.24	7.32	3.54
Pieza 5	3.07	3.66	4.68	5.51	8.11	7.68	5.78	7.32	6.59	6.49
MEDIA	4.3	4.36	5.57	7.05	8.42	9.38	7.75	6.63	6.05	4.78
RANGO	2.27	2.26	1.87	2.51	2.23	3.85	4.03	2.08	2.29	2.95
Sm	0.85	0.92	0.78	0.93	0.92	1.39	1.65	0.83	0.93	1.08

Tabla 7

$$\bar{x} = 6.43$$

$$\bar{R} = \frac{2.27 + 2.26 + 1.87 + 2.51 + 2.23 + 3.85 + 4.03 + 2.08 + 2.29 + 2.95}{10}$$

$$= 2.63$$

$$\bar{S}_M = \frac{0.85 + 0.92 + 0.78 + 0.93 + 0.92 + 1.39 + 1.65 + 0.83 + 0.93 + 1.08}{10}$$

$$= 1.03$$

A partir de estos estadísticos se pueden estimar la  $\sigma_{CP}$ :

$$\hat{\sigma}_{CP} = \frac{\bar{R}}{d_2(n=5)} = \frac{2.63}{2.3626} = 1.11$$

$$\hat{\sigma}_{CP} = \frac{\bar{S}_M}{c_4(n=5)} = \frac{1.03}{0.9400} = 1.11$$

Donde  $d_2$  y  $c_4$  son dos coeficientes tabulados, función del tamaño de la muestra, que corrigen el sesgo del estimador.

Si lo que se pretende es estimar la variabilidad total del proceso ( $\sigma_{LP}$ ) ha de hacerse a través de la desviación típica de todas las piezas

medidas con independencia de la muestra a la que pertenezcan. En los datos anteriores:

$$\sigma_{LP} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{50} (x_i - 6.43)^2}{50-1}} = 1$$

Como conclusión:

- Las dos estimaciones realizadas de la variabilidad inherente del proceso ( $\sigma_{CP}$ ) son bastante próximas (por ser el tamaño de la muestra pequeño,  $n = 5$ ).
- Como era de esperar, la estimación de la variabilidad total  $\sigma_{LP} > \sigma_{CP}$ .
- Los valores de los índices de capacidad son:
  - $C_p(\text{método } R) = 0.75$
  - $C_{pk}(\text{método } R) = 0.53$
  - $C_p(\text{método } S) = 0.76$
  - $C_{pk}(\text{método } S) = 0.55$
  - $P_p = 0.43$
  - $P_{pk} = 0.31$

Como ya se anticipó a la vista de la Figura 22, la variabilidad entre lotes es muy importante, lo que hace que la variabilidad a largo plazo sea muy superior a la variabilidad a corto plazo.

### 10.2.8 Construcción de un intervalo de confianza de los índices

Es preciso tener siempre presente que estos índices, al ser función de una variable aleatoria y estar calculados a través de estadísticos (valores numéricos calculados a partir de los datos muestrales), están sometidos a una variación aleatoria. Por esta razón la buena práctica estadística aconseja construir un intervalo de confianza. Este intervalo podría basarse en la conocida relación:

$$\sigma^2 \propto \frac{(n-1)s^2}{X_{n-1}^2}$$

que llevaría a los siguientes límites del intervalo de confianza para la varianza.

$$\frac{(n-1)s^2}{X_{n-1}^2(1-\alpha/2)} \leq \sigma^2 \leq \frac{(n-1)s^2}{X_{n-1}^2(\alpha/2)}$$

y para el índice

Sin embargo, para que esto sea cierto, es necesario una buena normalidad de los datos, hecho que no siempre se cumple. Por esta razón resulta más apropiado recurrir a métodos de estimación no paramétricos (por ejemplo, método de bootstrap).

#### 4. CASO DE PROCESOS NO NORMALES

##### 11.1 ¿ES LA NORMALIDAD LO “NORMAL”?

Existen procesos que incorporan a su naturaleza alguna causa especial que hace que su distribución no sea normal y consecuentemente no tiene sentido esperar su normalidad ni iniciar acciones correctoras para “meter el proceso en control estadístico”. Un caso típico son aquellos procesos que tienen alguna “barrera” que impide que la variable aleatoria pueda tomar valores desde  $-\infty$  a  $+\infty$ . Por ejemplo, un proceso de taladrado en el que el diámetro de la broca es el límite inferior del posible taladro, procesos químicos en el que se va añadiendo espesos de capas, etc.

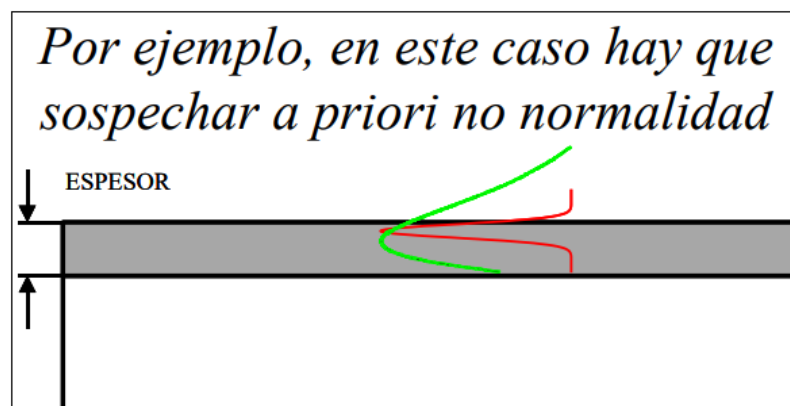


Figura 20: Si la variabilidad es grande en relación con el espesor medio, es de esperar una distribución asimétrica



## 11.2. CONSECUENCIAS QUE TIENE EN EL ANÁLISIS.

Cuando el proceso no sigue una distribución normal debido a causas intrínsecas a su propia naturaleza, lo expuesto anteriormente se ve afectado de la siguiente manera:

- a) La interpretación de los índices y en cuanto a la fracción de producto fuera de tolerancia ha de hacerse para la distribución que siga el proceso. Si no se conoce la distribución, utilizando la desigualdad de Tchebichef, sólo se puede asegurar que la cola situada más allá de  $3\sigma$  de la media no es superior a  $1/9=11\%$  (frente a  $0,135\%$  en el caso de un proceso normal).
- b) Los coeficientes utilizados para estimar  $\sigma_S$  y  $\sigma_R$  no son válidos y por lo tanto tampoco lo son los límites de control calculados para el gráfico de medias y recorridos.
- c) Es posible que la distribución del gráfico de medias sea a pesar de todo normal en virtud del TLC. No obstante, si la distribución del proceso tiene un grado elevado de asimetría sería necesario que el tamaño de la muestra fuese superior a 25 o 30 (en lugar de 4 o 5 como es habitual) para que se pudiera considerar normal a la distribución de las medias.

En estos casos, lo más recomendable es proceder como sigue:

- Puesto que la causa de no normalidad es intrínseca al proceso, se toma una muestra de al menos 25 o 30 unidades, de manera que la distribución de la muestra sea lo más próxima a la normal posible. De esta manera se puede mantener el control de las derivas en el proceso.
- Estudiar la distribución (calcular media, desviación típica y coeficiente de asimetría). En el caso de que la distribución no se encuentre contenida en el intervalo de tolerancias, ya se puede anticipar la incapacidad del proceso. Si se encuentra contenida en el intervalo de tolerancias de manera muy ajustada, es posible que el resto de las causas de variación presentes en la operación habitual del proceso haga que parte de la producción esté fuera de tolerancia.

- Realizar una transformación de normalización de los datos. En general resultan eficaces las transformaciones de BoxCox (ver 11.3).
- Una vez determinada la transformación numérica de normalización, si se codifican los valores tomados en las muestras, se puede tratar como los procesos vistos hasta ahora. Naturalmente para interpretar los valores de los límites de control, fracción fuera de tolerancia, etc. es preciso deshacer la transformación.

### 11.3 TÉCNICAS DE NORMALIZACIÓN DE LOS DATOS

La técnica de normalización de datos más eficaz es la transformación de Box-Cox. Esta transformación es la siguiente:

$$x^{(\lambda)} = \begin{cases} \frac{(x+m)^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \ln(x+m) & (\lambda = 0) \end{cases} \quad (\forall x > -m) \quad (m > 0)$$

Es continua en  $\lambda$ , puesto que:

$$\lim_{\lambda \rightarrow 0} x^{(\lambda)} = \ln(x+m)$$

La función de verosimilitud de los datos normalizados es:

$$\hat{\sigma}^2(\lambda) = \frac{1}{n} \sum_{i=1}^n (x^{(\lambda)} - \bar{x}^{(\lambda)})^2$$

La técnica consiste en hallar el valor de  $\lambda$  que maximiza la verosimilitud (ver Figura 21).

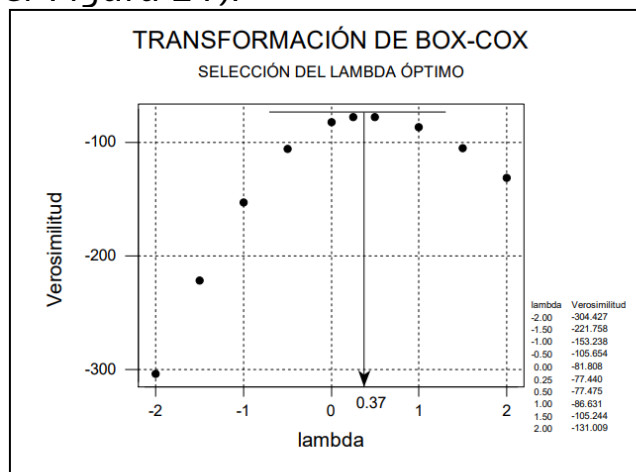


Figura 21: Transformación de Box-Cox para normalizar los datos

Según sea el origen de los datos, en algunos casos se puede esperar heterocedasticidad de estos (varianza no constante). Según esté relacionada la varianza y la media, resultan apropiadas las transformaciones indicadas en la Tabla 8.

Relación media-varianza	$\lambda$	Transformación estabilizadora de la varianza	Ejemplos típicos de aplicación
$\sigma \propto n^2$	-1	Inversa	
$\sigma \propto n^{1,5}$	-1/2	Inversa de la raíz cuadrada	
$\sigma \propto n$	0	Logaritmo	Análisis de varianzas muestrales, gráficos de control de series cortas en el que se conjuguen familias de varios espesores.
$\sigma \propto n^{0,5}$	1/2	Raíz cuadrada	Datos procedentes de una distribución de Poisson.
$\sigma \propto const$	1	No se transforma	

Tabla 8: Transformaciones de datos para eliminar la heterocedasticidad

### Ejemplo 5: Normalización de datos de un proceso no normal

En la Figura 25 se ha representado el histograma de un proceso de protección superficial. A la vista del citado histograma, se concluye la no normalidad de los datos. Se desea realizar una transformación que los normalice.

### APLICACIÓN MINITAB

Menú **Stat->Control** **Charts->Box-Cox**

**Transformation.** Los datos corresponden a muestras de tamaño 5 de 25 lotes. La respuesta proporcionada por MINITAB es la indicada en la Figura 24, en la que se indica  $\lambda=0.113$  como el valor óptimo. Una vez transformados los datos, se puede comprobar su normalidad mediante un histograma (ver Figura 25) o un papel probabilístico, de modo análogo al hecho anteriormente (ver Figura 26).

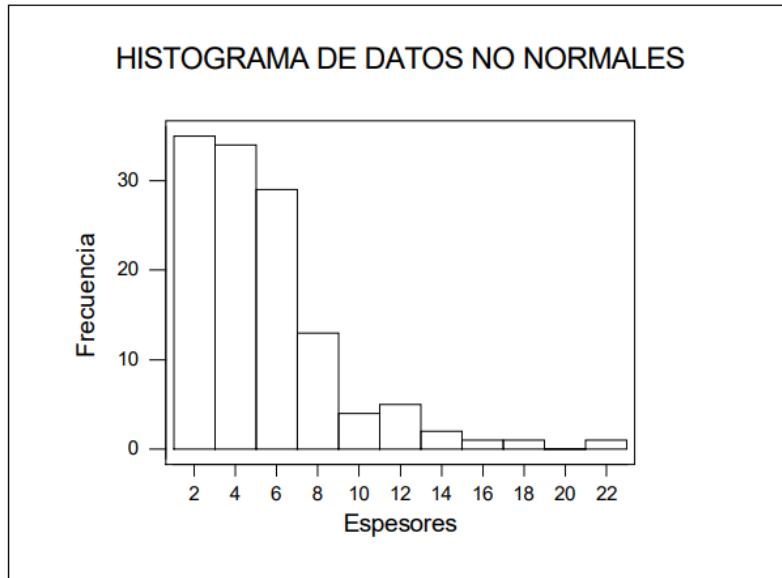


Figura 22 Histograma de datos no normales

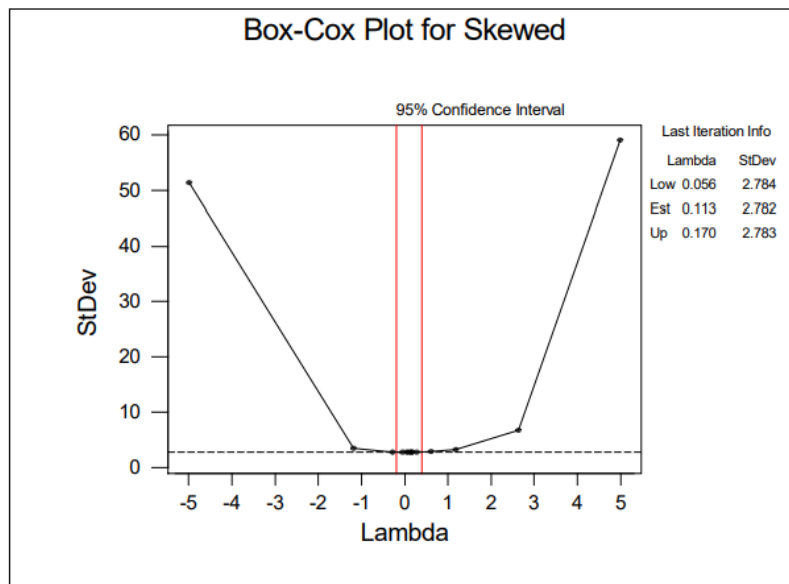


Figura 23 Histograma de datos de un proceso no normal

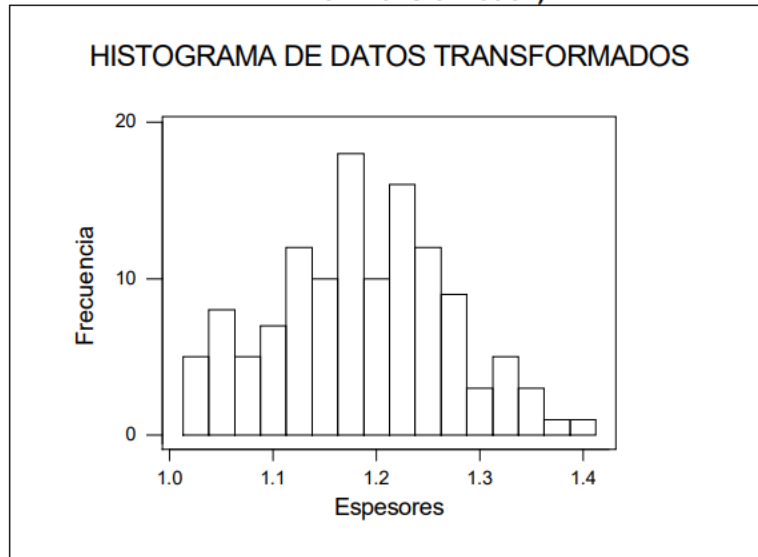


Figura 24: Transformación de Box- Cox (Stat->Control Charts -> Box-Cox Transformation)

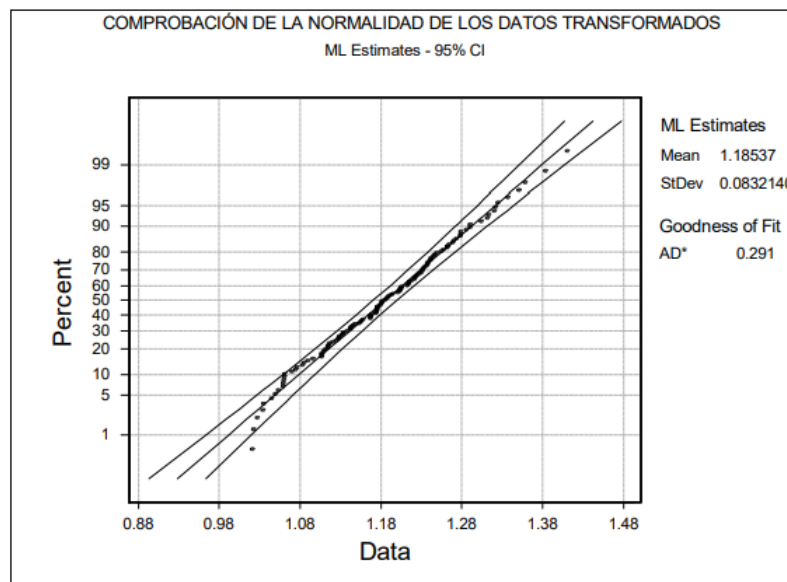


Figura 25: Histograma de los datos transformados (Graph-> Histogram)

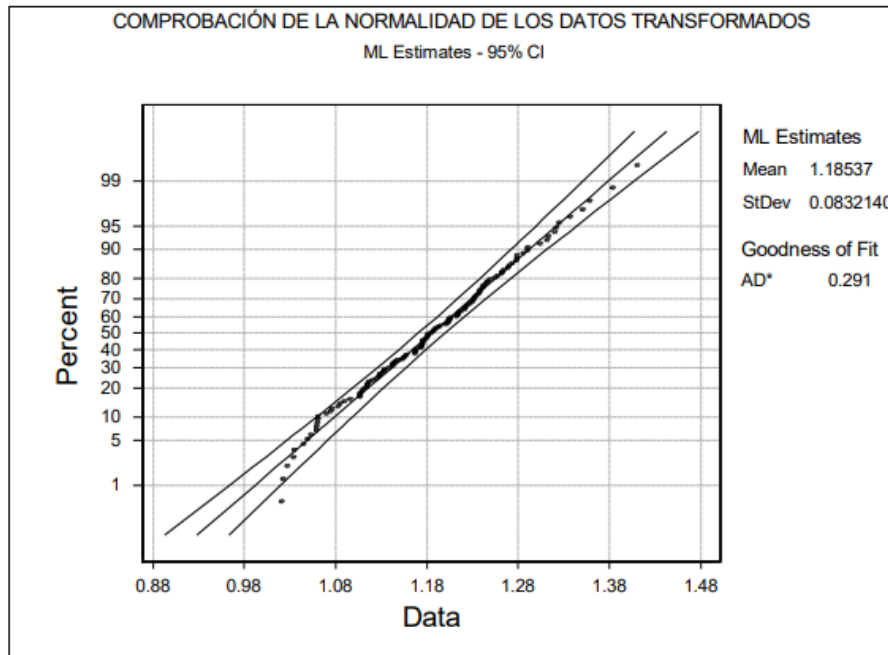


Figura 26: Comprobación de la normalidad de los datos transferidos

## 5. GRÁFICOS DE CONTROL PARA APLICACIONES ESPECIALES

### 5.1 GRÁFICOS PARA MUESTRAS UNITARIAS (I-MR)

En ocasiones interesa realizar gráficos de control a valores individuales, es decir, es la particularización a muestras de tamaño = 1. Al ser la muestra unitaria no resulta posible calcular la variabilidad dentro de la misma. Para resolver esto, se hace una agrupación artificial “de cada elemento con el siguiente”, de manera que se asocia la variación entre dos unidades consecutivas a la variación intermuestral. Esta variación se estima a través del “recorrido móvil” (moving range MR), que se calcula mediante la diferencia de dos valores consecutivos.

#### Ejemplo 6: Aplicación al análisis de datos de viscosidad de un proceso químico

Los datos proceden del libro “Implementing Six Sigma. Smarter Solution Using Statistical Methods” de Forrest W. Breyfogle III... Corresponden a la medida de la viscosidad en  $C_p$  de 20 lotes de un compuesto químico.

## APLICACIÓN MINITAB

Menú **Stat->Control Charts->I -MR**. La respuesta proporcionada por MINITAB es la indicada en la Figura 27.

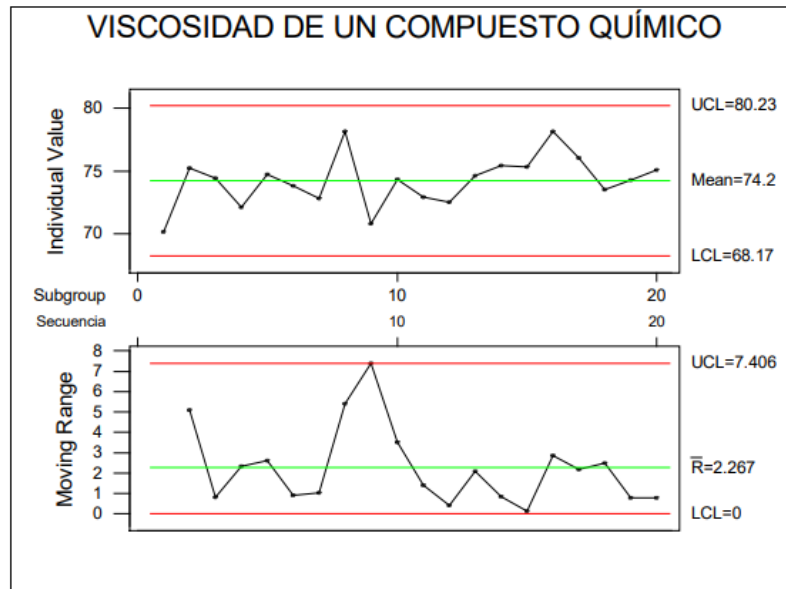


Figura 27: Gráfico I-MR de los valores de la viscosidad

### Referencias:

Fálco, A. (2006). Control estadístico de procesos. Universidad Pontificia Comillas, Madrid.