

Gráfico de Probabilidad Normal

Resumen

El *Gráfico de Probabilidad Normal* se usa para ayudar a juzgar si una muestra de datos numéricos proviene o no de una distribución normal. De no ser el caso, frecuentemente se puede determinar el tipo de alejamiento de la normalidad examinando la forma en la que los datos se desvían de la línea de referencia normal.

StatFolio de Ejemplo: *probplot.sgp*

Datos de Ejemplo:

El archivo *bottles.sf3* contiene la fuerza (*strength*) medida del rompimiento de $n = 100$ botellas de vidrio, similar al conjunto de datos contenido en Montgomery (2005). La tabla a continuación muestra una lista parcial de los datos de ese archivo:

<i>Strength</i> (Fuerza)
255
232
282
260
255
233
240
255
254
259
235
262

Ingreso de Datos

Los datos a analizar consisten de una sola columna numérica que contiene $n = 2$ o más observaciones.



- **Datos** : columna numérica que contiene los dato a resumir.
- **Selección**: selección de un subgrupo de datos.

Resumen del Análisis

El *Resumen del Análisis* muestra el número de observaciones en la columna de datos.

[Gráfico Normal de Probabilidad - strength](#)

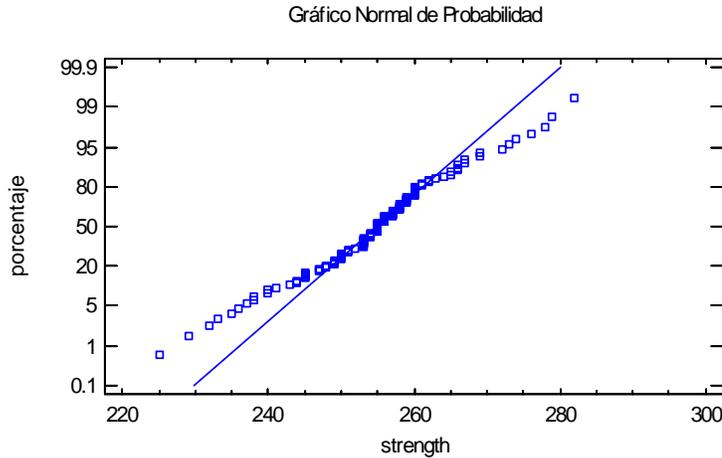
Datos/Variable: strength

100 valores con rango desde 225.0 a 282.0

También se presentan los valores máximo y mínimo.

Gráfico Normal de Probabilidad

Esta ventana exhibe el gráfico de probabilidad.



El gráfico se construye de la siguiente manera:

- Los datos se ordenan de menor a mayor y se determinan las estadísticas de orden. Por definición, la *j-ésima* estadística de orden es la *j-ésima* observación más pequeña en la muestra, denotada por $x_{(j)}$.
- Los datos se grafican en las posiciones

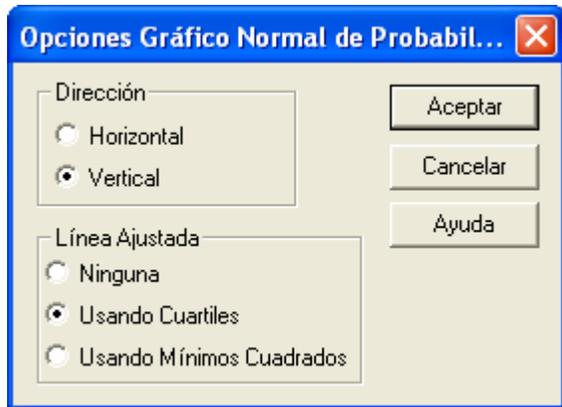
$$\left(x_{(j)}, \Phi^{-1}\left(\frac{j - 0.375}{n + 0.25}\right) \right) \quad (1)$$

donde $\Phi^{-1}(u)$ indica la distribución normal estándar evaluada en u .

- Si se desea, se ajusta una línea recta a los datos y se agrega al gráfico.

El gráfico de probabilidad normal se crea de tal forma que, si los datos son muestras aleatorias de una distribución normal, deben caer aproximadamente a lo largo de una línea recta. En el gráfico anterior, la desviación de los valores con respecto a la línea de referencia en ambos extremos indica que los datos pudieran provenir de una distribución con colas relativamente más largas que una distribución normal.

Opciones de Ventana

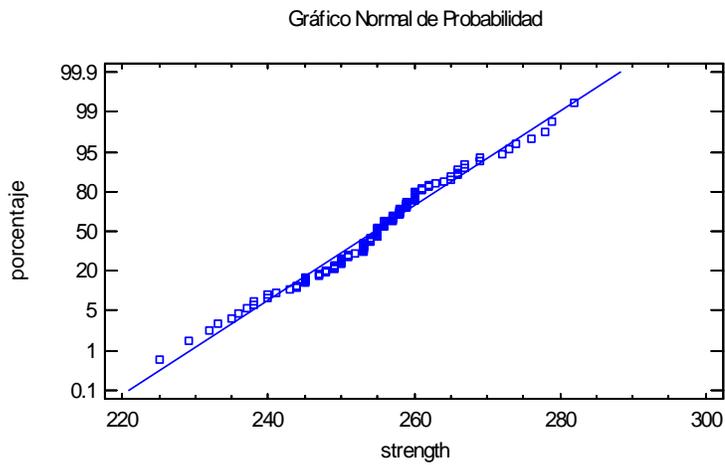


- **Dirección:** la orientación del gráfico. Si es *Vertical*, el *Porcentaje* se presenta en el eje vertical. Si es *Horizontal*, el *Porcentaje* se presenta en el eje horizontal.
- **Línea Ajustada:** el método usado para ajustar la línea de referencia a los datos. Si es *Usando Cuartiles*, la línea pasa por la mediana cuando el *Porcentaje* es de 50 con una pendiente determinada a partir del rango intercuartílico. Si es *Usando Mínimos Cuadrados*, la línea se ajusta con la regresión por mínimos cuadrados de los cuantiles normales de las estadísticas de orden. El primer método basado en los cuartiles da más peso a la forma de los datos cerca del centro y frecuentemente permite mostrar desviaciones de la normalidad en las colas que no serían evidentes usando el método de mínimos cuadrados.

Los valores por omisión de *Dirección* y *Línea Ajustada* se determinan a partir de la configuración en la pestaña *AED* de la caja de diálogo de las *Preferencias* en el menú *Editar*.

Ejemplo – Usando Mínimos Cuadrados

Poniendo *Tipo de Gráfico* en *Polígono* y marcando la caja *Acumulada* se obtiene una distribución acumulada de los datos:



Usando el método de *Mínimos Cuadrados* en los datos anteriores produce un gráfico con apariencia más “normal”, ya que la línea es forzada a ajustar tanto el centro como las colas.