

ANÁLISIS DE REGRESION

ANÁLISIS DE RESIDUALES

1. *GRAFICAS*
2. *ESTADISTICAS*

Gráfica de residuales para la medida

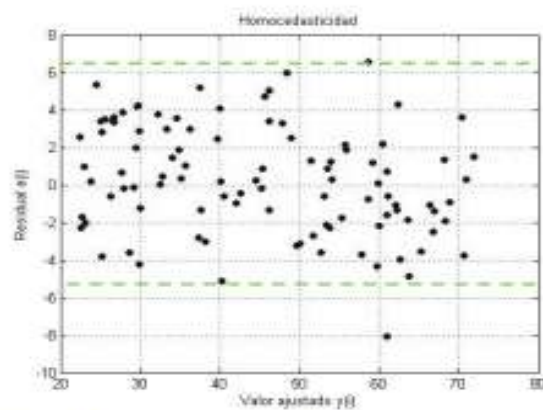
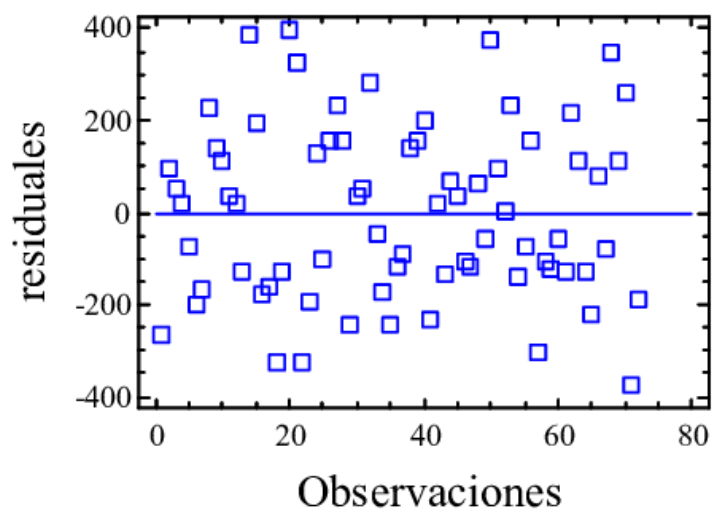


Figura 3. Residuales e_i vs. Valores ajustados \hat{y}_i

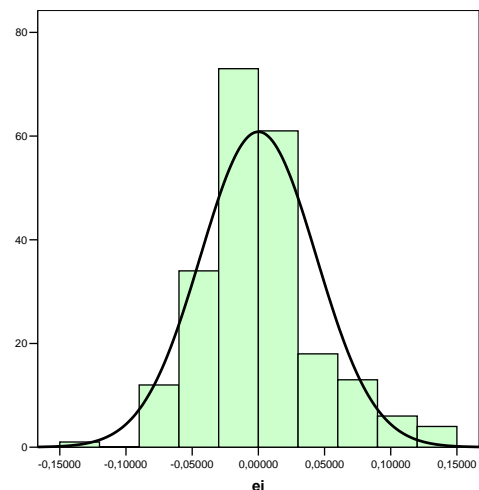
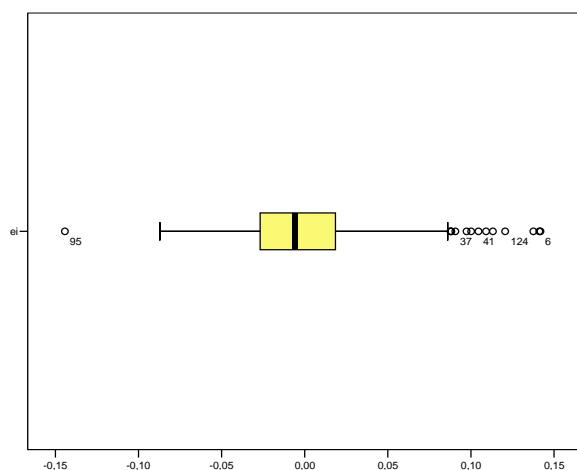
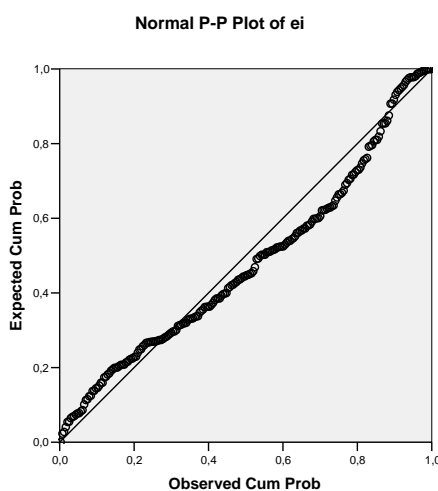
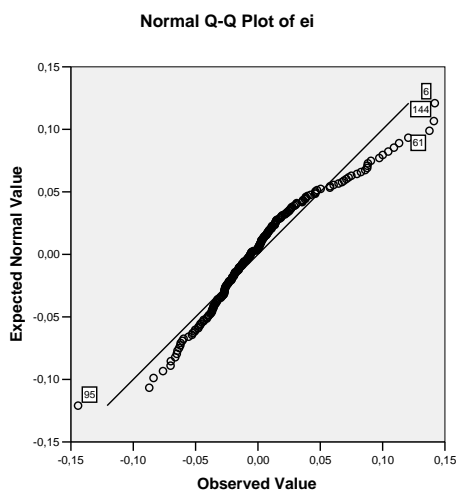
Gráficas de Residuales.

El análisis gráfico de los residuales es una forma muy efectiva de investigar la adecuación del ajuste de un modelo de regresión. Se recomienda graficar los residuales estudentizados porque tiene varianza constante.

Gráfica de Probabilidad Normal.

Los estadísticos t o F y los intervalos de confianza dependen de la suposición de normalidad. Existen una gran variedad de gráficos que permiten verificar si una muestra proviene de una población con distribución normal: histograma, tallos y hojas, boxplots, etc. Pero estos sólo son adecuados para muestras relativamente grandes.

Ejemplo: Comparación de gráficos para evaluar normalidad de las perturbaciones aleatorias



Otro método para comprobar la suposición de normalidad es trazar una gráfica de probabilidad normal de los residuales.

El gráfico de probabilidad normal es de gran ayuda para verificar el supuesto de que las perturbaciones aleatorias del modelo tienen distribución normal.

Para construir el **gráfico de probabilidad Normal** se procede a:

- Se ordenan los residuos en forma creciente: $e_{(1)} < e_{(2)} < \dots < e_{(n)}$
- Para cada valor "i" se calcula la probabilidad acumulada:

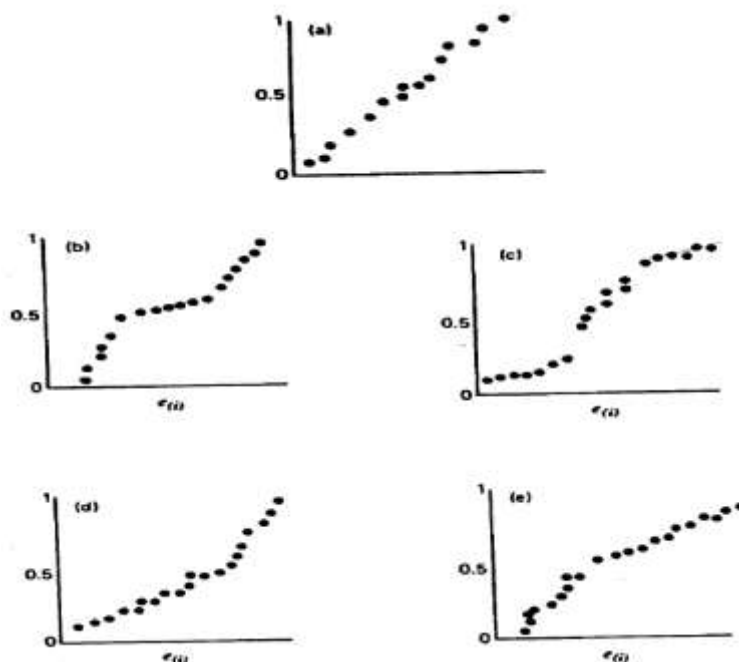
$$P_i = \frac{(i-0.5)}{n} \quad i = 1, 2, 3, \dots, n$$

- Graficar los pares $(e_{(i)}, P_i)$
- Si los puntos graficados se aproximan a una recta, podemos afirmar que las perturbaciones aleatorias del modelo tienen distribución normal.

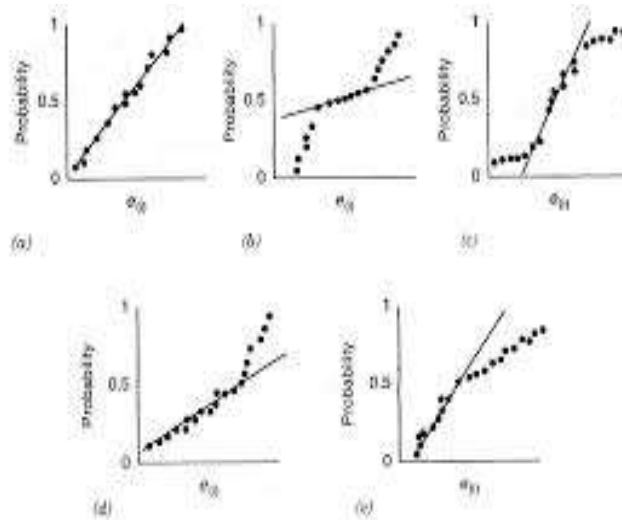
Es una gráfica diseñada para que al graficarse la distribución normal acumulada parezca una línea recta.

Ejemplo:

Comparación de gráficos para evaluar normalidad de las perturbaciones aleatorias



Identificando la normalidad



Analicemos las gráficas:

- a) Se puede ver una gráfica de probabilidad normal ideal.
- b) Distribución con colas que son demasiado pesadas.
- c) Distribución con colas más finas que la normal.
- d) Muestra que los datos provienen de una distribución con sesgo positivo, es decir con cola hacia la derecha, asimetría positiva.
- e) Los datos presentan un patrón correspondiente a una distribución con sesgo negativo, o sea, con cola hacia izquierda, asimetría negativa.

PRUEBAS ESTADÍSTICAS PARA LA NORMALIDAD

Entre las pruebas más conocidas se tiene:

Kolmogorov-Smirnov,

H_0 :

H_1 :

Shapiro –Wilk,

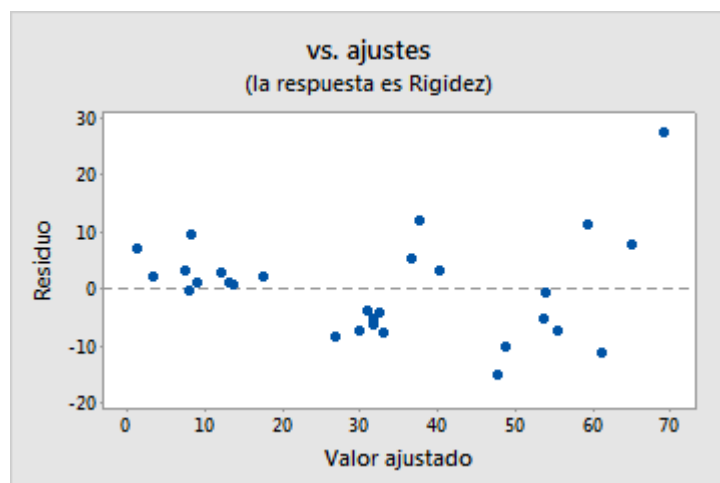
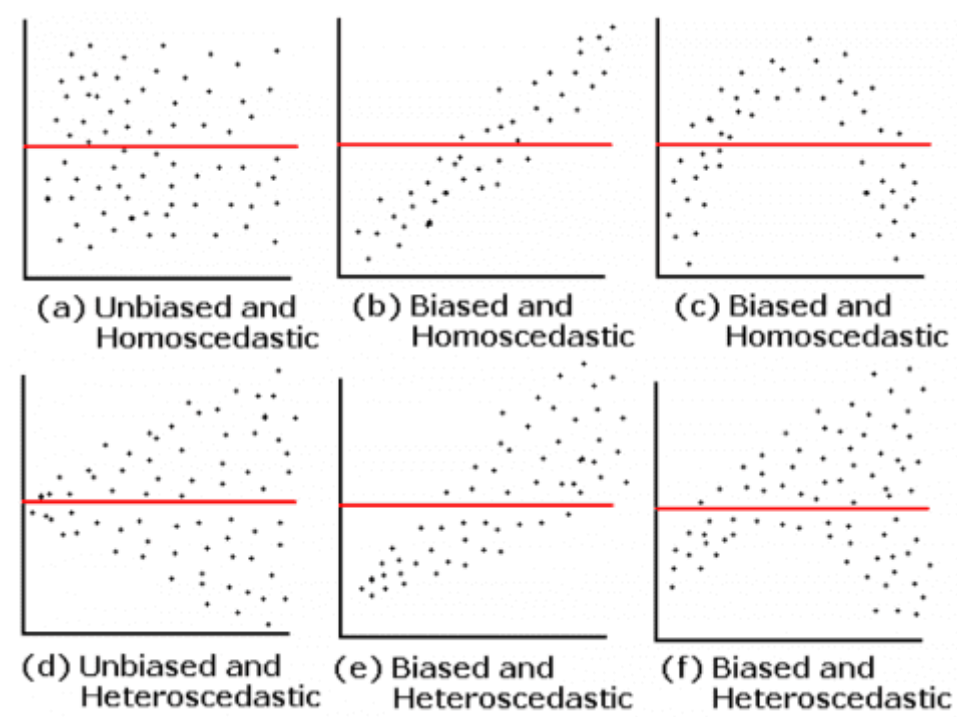
Mardia,

Bera y Jarque entre otros.

Gráfica de Residuales en Función de los Valores Ajustados \hat{y}_i

Se gráfica los residuales en función de los valores ajustados correspondientes \hat{y}_i . Es útil para detectar algunos tipos frecuentes de inadecuaciones del modelo.

Ejemplos de patrones:

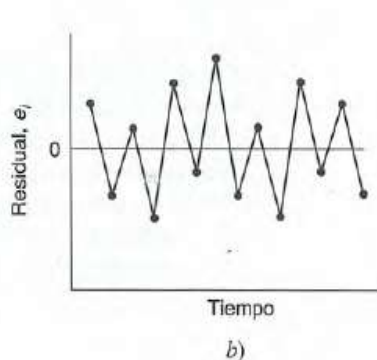
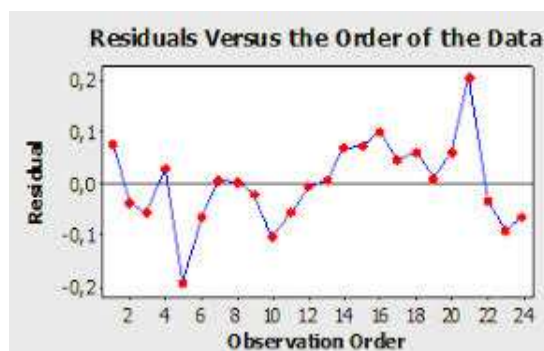
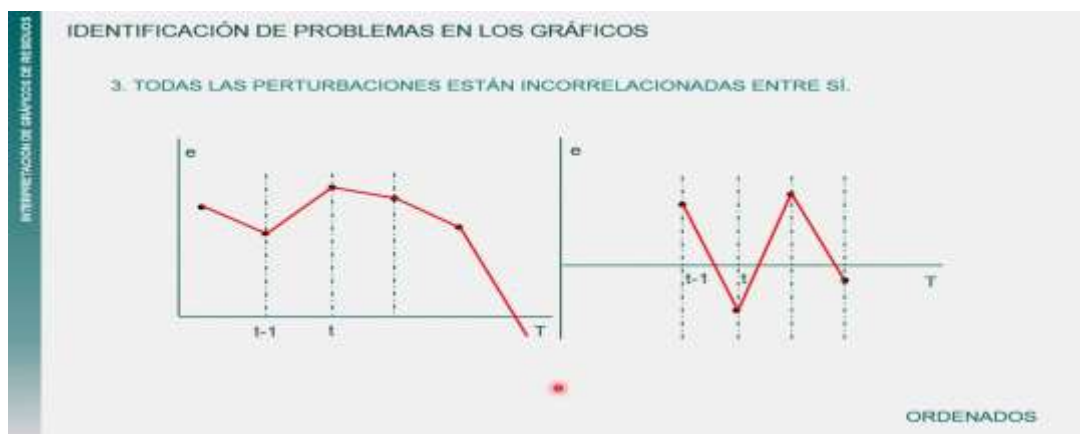


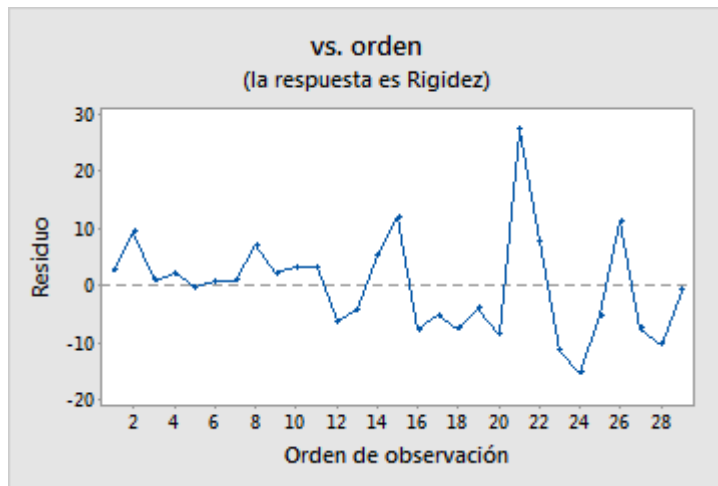
Las gráficas de residuos contra \hat{y}_i y las gráficas de probabilidad normal son útiles para identificar "outliers". Examinar residuos reescalados, como los estudentizados es una manera excelente para identificar observaciones atípicas potenciales.

ESTADÍSTICAS LEVENE....

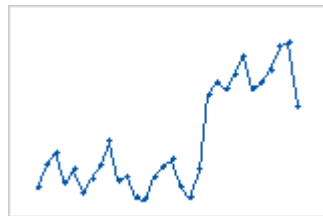
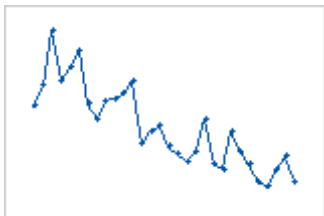
Gráfica de Residuales en el Tiempo.

Si se conoce la secuencia temporal de recolección de los datos se aconseja graficar los residuales en función de su orden en el tiempo. La gráfica de residuales en secuencial temporal puede indicar que los errores en un periodo se correlacionan con los de otros periodos. La correlación entre los errores del modelo en distintos periodos se llama autocorrelación. La presencia de autocorrelación es una violación potencialmente grave a las premisas básicas de la regresión.





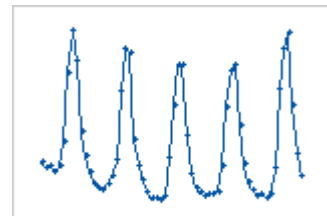
Si observa un patrón, investigue la causa. Los siguientes tipos de patrones pueden indicar que los residuos son dependientes.



Tendencia

Cambio

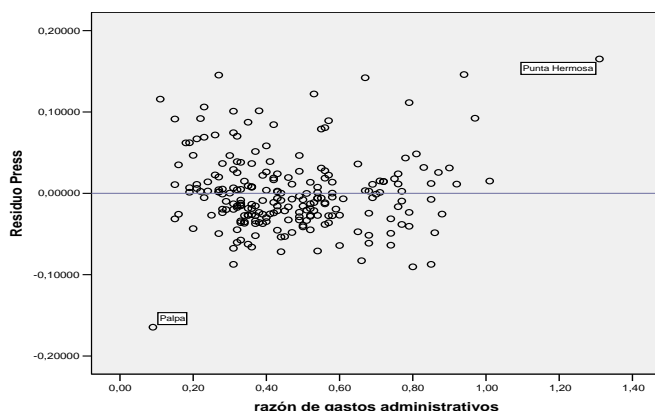
Ciclo

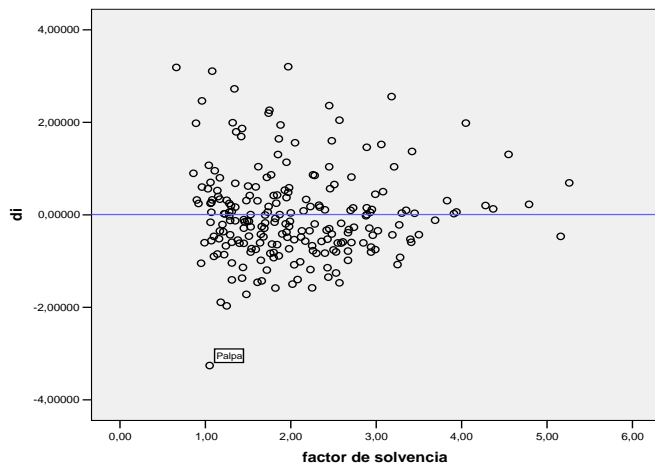


Gráfica de Residuales en Función del Regresor.

Una banda curva o una distribución no lineal implican en general que no es correcta la relación supuesta entre y y el regresor x_i por lo tanto deben considerarse algunas transformaciones.

Permite deducir si la existencia de heterocedasticidad o la falta de linealidad en el modelo son debidas a la variable explicativa representada.





Gráfica de Regresión Parcial.

Una gráfica de regresión parciales es una variación de la gráfica de residuales en función de predictor, que es una forma ampliada para estudiar la relación marginal de un regresor, dadas las demás variables· que están en el modelo. Esa gráfica puede ser muy útil para evaluar si se ha especificado en forma correcta la relación entre la respuesta y las variables regresoras

Las gráficas de regresión parcial consideran el papel marginal del regresor x_j dados los demás regresores que ya están en el modelo

El gráfico parcial de residuos, es útil para valorar la influencia real de una variable regresora, esto es, conocer la información nueva que aporta la variable regresora en estudio y que no aportan las otras variables.

Se deben hacer los gráficos entre variables regresoras

