

10. Transformación de Box-Cox

Otros factores que también pueden alterar la “estacionariedad”, así como la *linealidad* subyacente a una serie observada, son cambios muy abruptos en la misma. En otras palabras, la variabilidad en los datos es muy caótica en cierto intervalo. Esto se traduce en un rompimiento al requisito de estacionariedad débil, e.g. $\gamma_X(h)$ ya no únicamente depende de la distancia de rezago h sino también del punto en el tiempo y por lo tanto la propiedad (iii) de la Definición 5 no se cumple.

Una manera de suavizar dichos cambios es a través de su *transformación logarítmica*

$$z_t = \ln(x_t),$$

la cual tiende a suprimir fluctuaciones grandes. Otra posible transformación esta dada mediante

$$y_t = \begin{cases} (x_t^\lambda - 1)/\lambda, & \lambda \neq 0 \\ \ln x_t, & \lambda = 0. \end{cases} \quad (28)$$

Esta transformación, conocida como *la transformación Box-Cox*, también es usada para aproximar a normalidad.

11. Residuales

El objetivo de las transformaciones a estacionariedad, es precisamente el de obtener una serie, la cual no exhiba una tendencia o ciclicidad aparente con la finalidad de poder modelar este componente de manera independiente. Sin embargo, una vez que ejecutamos algunas de las técnicas descritas anteriormente existen una serie de pruebas que se deben de realizar a los *residuales*.

11.1. Dependencia

La primera de estas pruebas, se refiere a la independencia entre los elementos de la serie dada por los residuales, $\{Y_t\}$. Por ejemplo, si los datos fuesen iid, entonces el problema se resolvería fácilmente mediante el uso de algunas características como la media y la varianza subyacente a los residuales. Sin embargo, cuando se analizan series de tiempo,

normalmente existe dependencia y se requiere de modelos más elaborados como los que veremos en lo que resta del curso.

Una manera de visualizar la dependencia en una serie (e.g. los residuales) es a través de examinar gráficamente las *densidades* correspondientes a $\{X_t, X_{t+1}\}$, $\{X_t, X_{t+2}\}, \dots, \{X_t, X_{t+h}\}$ y ver si estas exhiben una dependencia evidente. En la Figura (6) se observa un ejemplo de estas densidades: tanto la correspondiente a los datos como la correspondiente a un modelo dado. El dibujo en puntos representa la estimación de la superficie dados los datos $\{x_1, x_2, \dots, x_{n-1}\}$ vs. $\{x_2, x_3, \dots, x_n\}$. Sin embargo, un diagrama de dispersion (“scatterplot”).

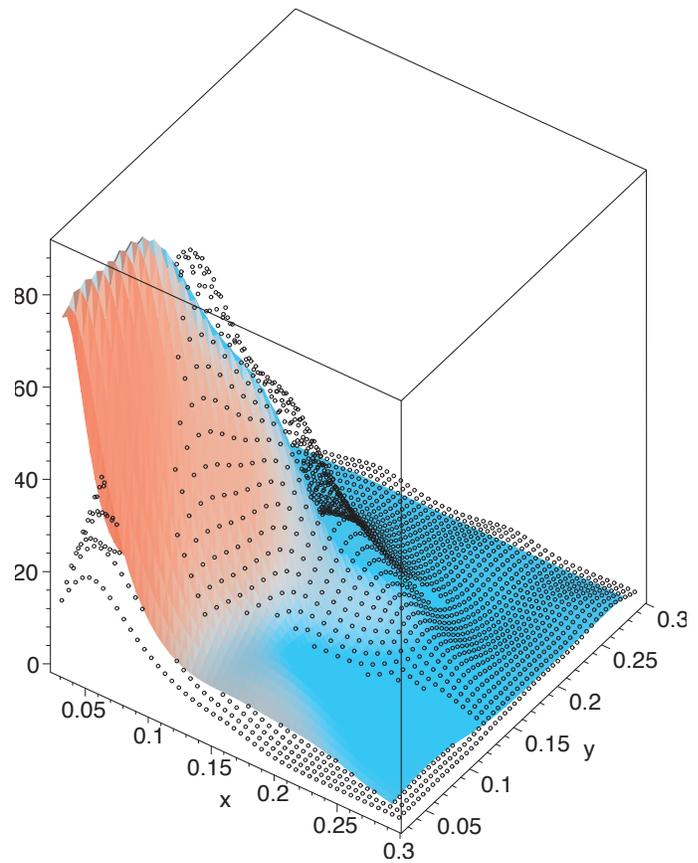


Figura 6: La densidad conjunta de $\{X_t = x, X_{t+1} = y\}$. El dibujo en puntos denota los datos observados, y el dibujo en “solido” un modelo (no-lineal) ajustado.

Otra forma, también gráfica, pero más limitada es mediante el gráfico de la ACF muestral (también conocido como *correlograma*), ver Figura 3. Esta gráfica, es claramente menos robusta ya que únicamente se considera la correlación entre x_t y x_{t+h} , $h = 1, 2, \dots$. Sin embargo, bajo el supuesto de

estacionariedad débil, que domina los modelos más comunes (“clásicos”) y que veremos en el siguiente capítulo, este método es aceptable.

En teoría, para una muestra grande (n grande) la ACF correspondientes a una sucesión Y_1, \dots, Y_n de variables iid y con varianza finita se distribuyen $N(0, 1/n)$ aproximadamente. Por lo tanto si y_1, \dots, y_n es una realización de una sucesión iid, entonces alrededor del 95 % de las ACF muestrales deberán estar contenidas en el intervalo formado por $\pm 1.96/\sqrt{n}$. Muchas librerías de lenguajes de programación ya calculan y muestran dichos intervalos de confianza. Ver Figura 3.

En vez de checar si las autocorrelaciones caen dentro del intervalo, también se puede considerar la estadística dada por

$$Q = n \sum_{j=1}^h \hat{\rho}^2(j). \quad (29)$$

También conocida como la **prueba de portmanteau**. Por el mismo resultado enunciado anteriormente, $Q \sim \chi^2(h)$. Por lo tanto se rechaza la hipótesis de iid al nivel α si

$$Q > \chi_{1-\alpha}^2(h).$$

Algunos programas, en particular ITSM, utilizan una mejor aproximación a la distribución $\chi^2(h)$, conocida como la **prueba de Ljung-Box** y dada por la estadística

$$Q = n(n+2) \sum_{j=1}^h \hat{\rho}^2(j)/(n-j). \quad (30)$$

11.2. Normalidad

Otra característica que suele solicitarse, al menos en el enfoque clásico, es que los residuales sigan una distribución normal.

Definición 6. Sea $Z = (Z_1, Z_2, \dots, Z_n)$ un vector aleatorio. Se dice que Z se distribuye normalmente, $Z \sim N_n(\mu, \Sigma)$, si

$$N_n(\mathbf{z}; \mu, \Sigma) = |2\pi\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2}(\mathbf{z} - \mu)' \Sigma^{-1}(\mathbf{z} - \mu) \right\},$$

donde $\mathbf{z} := (z_1, \dots, z_n)'$, $\mu := (\mu_1, \dots, \mu_n)'$ y

$$\Sigma := \{\sigma_{ij}; i, j = 1, \dots, n\}$$

denota una matriz simétrica y positiva definida, conocida como la matriz de covarianzas σ_{ij} .

Sean $Y_{(1)} < Y_{(2)} < \dots < Y_{(n)}$ las estadísticas de orden de una muestra aleatoria de la distribución $N(\mu, \sigma^2)$. Si $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ son las estadísticas de orden de una muestra aleatoria de la distribución $N(0, 1)$ entonces

$$\mathbb{E}[Y_{(j)}] = \mu + \sigma m_j,$$

donde $m_j = \mathbb{E}[X_{(j)}]$, $j = 1, \dots, n$. A la gráfica de dispersión de los puntos $(m_1, Y_{(1)}), \dots, (m_n, Y_{(n)})$ se le conoce como **qq-plot Gaussiana**. Si la suposición de normalidad es correcta, la gráfica de qq-plot Gaussiana debe ser aproximadamente lineal. Consecuentemente, el cuadrado de la correlación entre los puntos $(m_i, Y_{(i)})$, $i = 1, \dots, n$ debe de ser cercana a uno. Entonces, el supuesto de normalidad se rechaza si el cuadrado del coeficiente de correlación R^2 es suficientemente pequeño. Si aproximamos m_i mediante $\Phi^{-1}((i-0.5)/n)$ entonces

$$R^2 = \frac{[\sum_{i=1}^n (Y_{(i)} - \bar{Y}) \Phi^{-1}(\frac{i-0.5}{n})]^2}{\sum_{i=1}^n (Y_{(i)} - \bar{Y})^2 \sum_{i=1}^n (\Phi^{-1}(\frac{i-0.5}{n}))^2} \quad (31)$$

12. Enfoque general en el estudio de series de tiempo

- Dibujar la serie y examinar las principales características de la serie, en particular identificar posibles
 - (a) Componentes de tendencias
 - (b) Componentes de ciclicidad
 - (c) Cambios abruptos y/o observaciones aberrantes.
- Hacer las transformaciones necesarias para que los residuales sean lo más estacionarios posibles.
- Hacer algunas pruebas de independencia (y de ser el caso normalidad) a los residuales resultantes.
- Escoger el modelo que mejor se ajuste a los residuales.
- Predicción será posible mediante la predicción de de los residuales e invirtiendo las transformaciones que se efectuaron a la serie original.