Series de tiempo para datos de valor entero o para datos categóricos

Julieta Ruiz, Sofía Bello, Andrés García

druizca@unal.edu.co
abellod@unal.edu.co
andgarcialo@unal.edu.co

Departamento de Estadistica
Universidad Nacional de Colombia

Noviembre 2023

Agenda

UNIVERSIDAD NACIONAL DE COLOMBIA

Motivación

En la cotidianidad es posible encontrar fenómenos en los cuales las características de interés son tomadas en valores enteros o categóricos. Por ejemplo, el número de muertes en una ciudad, el número de nacimientos o.

Por dicha razón se hace necesario la utilización de modelos que reflejen la naturaleza de estos datos, teniendo en cuenta el tipo de dato que se tiene es posible encontrar dos tipos de categorías.

1. Series de tiempo de conteo.
2. Series de tiempo categóricas.

A continuación mencionaremos la construcción y principales características de los modelos utilizados para analizar este tipo de series de tiempo.

Condiciones de estacionalidad

Las condiciones de estacionalidad son análogas a las utilizadas para las series de tiempo para datos continuos, teniendo en cuenta que las distribuciones utilizadas serán discretas.

- Sea $x_t$un proceso estocástico, se dice que este es estacionario en el sentido débil si:
  - $E|X_t|^2 < \infty$ para todo t.
  - $\mu_t = E[X_t] = m$ para todo t.
  - $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ para todo r,s y t.

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelos DARMA

Los modelos DARMA buscan reproducir las características de versatilidad, generalidad y disponibilidad de los modelos ARMA clásicos, pero reemplazando la combinación lineal del caso de los valores continuos por el modelo lineal probabilístico.

Definiciones

Se necesitan algunas definiciones para definir los procesos, estos se muestran a continuación:

- Sea $\{Y_t\}$ una sucesión de v.a iid que toman valores en un conjunto $E \in \mathbb{Z}$ con $P(Y_t = i) = \pi(i)$ para todo $i \in E$.
- Sean $\{U_t\}$ y $\{V_t\}$ sucesiones independientes de variables aleatorias binarias iid tales que $P(U_t = 1) = \beta$ y $P(V_t = 1) = \tau$ con $\beta, \tau \in [0, 1]$.
- Sean $\{D_t\}$ y $\{A_t\}$ sucesiones de v.a iid que toman valores $0, 1, 2, \cdots, q$ para un valor fijo q en los enteros positivos.

Proceso autorregresivo discreto de orden 1 DAR(1)

El modelo DAR(1) definido por:

$$Z_t = V_t Z_{t-1} + (1 - V_t) Y_t$$

Es decir que $Z_t = Z_{t-1}$ con probabilidad τ y $Z_t = Y_t$ con probabilidad $(1 - \tau)$, además la probabilidad de transición viene dada por:

$$P(Z_t = i | Z_{t-1} = k) = \tau \delta_{ki} + (1 - \tau)\pi_i$$

Donde $\pi_i = P(Y_t = i)$ y δ_{ki} es la función delta de Kronecher, además se tiene que $corr(Z_t, Z_{t+h}) = \tau^h$

UNIVERSIDAD NACIONAL DE COLOMBIA

Proceso autorregresivo discreto de orden p DAR(p)

El modelo general DAR(p) viene dado por:

$$Z_t = V_t Z_{t-A_t} + (1 - V_t) Y_t$$

Para $t = -q, -q+1, \cdots$, lo cual quiere decir que con probabilidad τ, Z_t es alguno de los anteriores valores $Z_{t-1}, \cdots, Z_{t-p}$ o de otro modo es Y_t, de modo general se tiene que $\rho(h) = \tau^h$

DMA(1)

El modelo DMA(1) de $\{s_t\}$ satisface:

$$X_t = V_t Y_t + (1 - V_t) Y_{t-1}$$

Es decir que $X_t = Y_t$ con probabilidad τ y $Z_t = Y_{t-1}$ con probabilidad $1 - \tau$. Además, se tiene que:

$$\rho(h) = \begin{cases} \tau(1-\tau) & \text{para } h = 1 \\ 0 & \text{para } h = 2, 3, \cdots \end{cases}$$

DARMA(p,q+1)

Se obtiene de acuerdo a:

$$S_t = U_t Y_{t-D_t} + (1-U_t)Z_{t-(q+1)}, t = 1, 2, \cdots$$
$$Z_t = V_t Z_{t-A_t} + (1-v_t)Y_t, t = -q, -q+1, \cdots$$

Z_t es un proceso DAR(p). Lo anterior nos indica que con probabilidad β, S_t es uno de los valores $Y_t, \cdots, Y_{t-q}$ o de otro modo es igual a la cola autorregresiva $Z_{t-(q+1)}$.

- si $\beta = 0$, entonces, $S_t = Z_{t-(q+1)}$ y, por tanto, es un proceso DAR(p)
- si $\beta = 1$, entonces $S_t = Y_{t-D_t}$ y, por tanto, es un proceso DMA(q)

Jacob y Lewis (1978) muestran que existe una distribución estacionaria para el vector aleatorio $(Z_{-q-p}, \cdots, Z_{q-1})$. Si el proceso DAR(p) se inicia con esa distribución estacionaria, el proceso DARMA $\{S_t : t \in \mathbb{N}\}$ es entonces un proceso estacionario con distribución marginal π.

UNIVERSIDAD NACIONAL DE COLOMBIA

Desventajas

Dicha analogía mencionada se tiene de una manera artificial y solo desde el punto de vista teórico. En muchos casos el comportamiento de estos modelos no es el mejor, por ejemplo:

- El proceso DAR(1) no produce series correlacionadas negativamente, por lo tanto, no puede modelar cualquier tipo de serie de tiempo de variables discretas.
- En el proceso DARMA(1,q+1), $Z_{t-(q+1)}$ tiene toda la información del pasado del proceso, teniendo en cuenta que Y_{t-D_t} es independiente del histórico, si este se elige (el cual tiene probabilidad β), se pierde toda la información histórica del proceso.

Operador $'\circ'$ (Binomial Thinning)

La recursión clásica del modelo AR(1), $Y_t = \alpha \dot{Y}_{t-1} + \epsilon_t$ no puede ser aplicada a procesos de conteos, incluso si se asume que ϵ_t es un valor entero. Puede ocurrir que Y_t no sea un valor entero, es decir, que $\alpha\cdot$ no mantenga el rango discreto.
Sea $\{Y_i, i \in \mathbb{N}\}$ una sucesión de variables aleatorias Bernoulli iid con media α, independientes de X, una variable aleatoria de valores enteros no negativos. El operador $'\circ'$ de Steutel y Van Harn está definido por:

$$\alpha \circ X = \begin{cases} \sum_{i=1}^{X} Y_i & \text{para } X > 0 \\ 0 & \text{para } X = 0 \end{cases}$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades del Operador $'\circ'$

Sean X y Y variables aleatorias de valores enteros no negativos y sean $\alpha \in [0,1]$ y $\beta \in [0,1]$, entonces se tiene que:

1. $0 \circ X = 0$
2. $1 \circ X = X$
3. $E[\alpha \circ X] = \alpha E[X]$
4. $E[(\alpha \circ X)^2] = \alpha(1-\alpha)E[X] + \alpha^2 E[(X)^2]$
5. $\alpha \circ (X + Y) = \alpha \circ X + \alpha \circ Y$
6. $E[(\alpha \circ X - \alpha \circ Y)^2] = \alpha(1-\alpha)E[|X - Y|] + \alpha^2 E[(X - Y)^2]$
7. $\beta \circ (\alpha \circ X) \stackrel{d}{=} (\beta\alpha) \circ X$
8. Si las variables y_i incluidas en $\alpha \circ X$ son independientes de las variables Y_j incluidas en $\beta \circ\ y$, entonces. $E[(\alpha \circ X)(\beta \circ Y)] = \alpha\beta E[XY]$

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelo INAR(1)

El proceso autorregresivo de valores enteros no negativos de primer orden, INAR(1), es eficaz en proceso, en donde en el tiempo t se encuentran los sobrevivientes en el tiempo, $t-1$ más un proceso de innovación no negativo. El proceso INAR(1), $\{X_t, t = \pm 1, \pm 2, \pm 3, \cdots\}$, tiene la siguiente representación:

$$X_t = \alpha \circ X_{t-1} + \epsilon_t$$

Con ϵ_t, $t \in \mathbb{Z}$ es una variable aleatoria (i.i.d) de media μ_ϵ y varianza σ_ϵ^2 independiente de X_{t-1}.

Interpretación

Los componentes del proceso INAR(1) en el tiempo t, X_t son:

- Los sobrevivientes de los elementos del proceso en el tiempo $t-1$, X_{t-1} cada uno con probabilidad de sobrevivencia α ($\alpha \circ X_{t-1}$).
- Los elementos que entraron al sistema en el intervalo $(t-1, t]$ como término de innovación (ϵ_t).

Propiedades generales

Suponiendo que $E\left[\epsilon_t\right] = \mu_\epsilon < \infty$, $V\left[\epsilon_t\right] = \sigma_\epsilon^2 < \infty$ y $\alpha \in [0, 1)$, entonces el proceso INAR(1) es estacionario y se tiene que:

Valor esperado: $E\left(X_t\right) = \frac{\mu_\epsilon}{(1-\alpha)}$

Varianza: $V\left(X_t\right) = \gamma(0) = \frac{\mu_\epsilon \alpha + \sigma_\epsilon^2}{1-\alpha^2}$

Función de autocovarianza:
$\gamma(h) = \mathrm{Cov}\left(X_{t-h}, X_t\right) = \alpha^h V\left(X_{t-h}\right) = \alpha^h \gamma(0)$

Función de autocorrelación:
$\rho(k) = Corr(X_t, X_{t-h}) = \alpha^k, k = 1, 2, 3, \ldots$

$\rho(k)$ presenta un decaimiento exponencial con retardo k y este nunca se anula.

Distribución marginal.

Utilizando la función generadora de probabilidad de X_t, denotada por G_{X_t}, junto con algunas propiedades del operador " $\circ$ ", es posible verificar que la distribución marginal del modelo se puede expresar como:

$$X_t \stackrel{d}{=} \sum_{i=0}^{\infty} \alpha^i \circ \epsilon_{t-i}$$

La cual está determinada por la distribución de ϵ_t, si ϵ_t tiene distribución de Poisson con media λ entonces X_t también tendrá distribución de Poisson. De este modo, la probabilidad de transición viene dada por:

$$\begin{aligned} p_{ij} &= P(X_t = j | X_{t-1} = i) \\ &= e^{-\lambda} \sum_{k=0}^{min(i,j)} \binom{i}{j} \alpha^k (1-\alpha)^{i-k} \lambda^{j-k} / (j-k)! \end{aligned}$$

Donde:
$\mu_\epsilon = \sigma^2_\epsilon = \lambda$, con lo cual $X_t \sim Poiss\left(\frac{\lambda}{1-\alpha}\right)$ y $\epsilon_t \sim Poiss(\lambda)$.

Estimaciones

Estimadores de Yule-Walker

Teniendo en cuenta que $\widehat{\alpha} = \widehat{\rho(1)} := \widehat{\gamma(1)}/\widehat{\gamma(0)}$, de este modo se tiene para $k = 1$, que:

$$\widehat{\alpha} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-1} \left(X_t - \bar{X}_t\right)\left(X_{t+1} - \bar{X}_t\right)}{\sum_{t=1}^{n} \left(X_t - \bar{X}_t\right)^2}$$

Donde $\bar{X}_t$ es la media muestral.

Si se ajusta un Poisson INAR(1) tenemos otro parametro λ (media de las innovaciones). Teniendo en cuenta que $\epsilon_t = X_t - \alpha \circ X_{t-1}$, entonces tenemos que:

$$
\begin{aligned}
\widehat{E[\epsilon_t]} = \widehat{\lambda} &= \widehat{E[X_t]} - \widehat{\alpha}\widehat{E[X_{t-1}]} \\
&= \sum_{t=1}^{n} \frac{X_t}{n} - \widehat{\alpha} \sum_{t=1}^{n} X_{t-1} \\
&= \frac{1}{n} \left[\sum_{t=1}^{n} (X_t - \widehat{\alpha} X_{t-1}) \right]
\end{aligned}
$$

Estimaciones

Estimadores de Mínimos Cuadrados Condicionales

Se basa en la minimización de la suma de desviaciones cuadradas acerca de la esperanza condicional, teniendo en cuenta la media condicional de X_t dado X_{t-1}:

$$E[X_t|X_{t-1}] = \alpha X_{t-1} + \lambda$$

los estimadores de mínimos cuadrados condicionales para α y λ son los valores que minimizan a:

$$Q_n(\alpha, \lambda) = \sum_{t=1}^{n} \left(X_t - E\left[X_t / X_{t-1}\right]\right)^2$$

En este orden de ideas se deriva con respecto α y λ e igualando a 0 se obtiene que los estimadores son dados por:

UNIVERSIDAD NACIONAL DE COLOMBIA

Estimaciones

Estimadores de Mínimos Cuadrados Condicionales

$$\begin{aligned}
Q_n(\alpha,\lambda) &= \sum_{t=1}^{n}(X_t - E[X_t|X_{t-1}])^2 = \sum_{t=1}^{n}(X_t - (\alpha X_{t-1}+\lambda))^2 \\
&= \sum_{t=1}^{n}\left(X_t^2 - 2X_t(\alpha X_{t-1}+\lambda) + (\alpha X_{t-1}+\lambda)^2\right) \\
&= \sum_{t=1}^{n}(X_t^2 - 2\alpha X_t X_{t-1} - 2\lambda X_t + \alpha^2 X_{t-1}^2 + 2\alpha\lambda X_{t-1} + \lambda^2)
\end{aligned}$$

Derivando $Q_n(\alpha,\lambda)$ con respecto a α y λ, se tiene que:

$$\frac{\partial Q_n(\alpha,\lambda)}{\partial\alpha} = \sum_{t=1}^{n}\left(-2X_tX_{t-1} + 2\alpha X_{t-1}^2 + 2\lambda X_{t-1}\right) = 0 \qquad (1)$$

$$\frac{\partial Q_n(\alpha,\lambda)}{\partial\lambda} = \sum_{t=1}^{n}(-2X_t + 2\alpha X_{t-1} + 2\lambda) = 0 \qquad (2)$$

De las anteriores derivadas de $Q_n(\alpha, \lambda)$, se obtiene que $\hat{\lambda}$, está dado por:

$$\widehat{\lambda} = \frac{1}{n} \sum_{t=1}^{n} (X_t - \widehat{\alpha} X_{t-1})$$

Sustituyendo el estimador de λ en la ecuación (1) y despejando con respecto a α, obtenemos la estimación para α, dada por:

$$\widehat{\alpha} = \frac{n \sum_{t=1}^{n} X_t X_{t-1} - \sum_{t=1}^{n} X_t \sum_{t=1}^{n} X_{t-1}}{n \sum_{t=1}^{n} X_{t-1}^2 - \left(\sum_{t=1}^{n} X_{t-1} \right)^2}$$

Modelo INAR(p)

Sea X_t con $t = \pm 1, \pm 2, \pm 3, ...$ una sucesión de variables aleatorias de valores enteros no negativos, tales que:

$$X_t = \sum_{i=1}^{p} \alpha_i \circ X_{t-i} + \epsilon_t$$
$$= \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \cdots + \alpha_p \circ X_{t-p} + \epsilon_t$$

Donde se tiene que:

- $\{\epsilon_t : t = 0, \pm 1, \pm 2, \cdots\}$ es una sucesión de v.a iid con media $\mu_\epsilon < \infty$ y varianza $\sigma_\epsilon^2 < \infty$.
- α_i con $i = 1, 2, \cdots, p$ son constantes no negativas tales que $\sum_{i=1}^{p} \alpha_i < 1$.
- La distribución del vector $[(\alpha_1 \circ X_t, \alpha_2 \circ X_t, \cdots, \alpha_p \circ X_t | X_t = x)]$ es multinomial con parámetros $(x, \alpha_1, \cdots, \alpha_p)$ y es independiente de la historia pasada del proceso.

Propiedades

- **Valor esperado:**

$$E[X_t] = \frac{\mu_\epsilon}{1 - \sum_{i=1}^{p} \alpha_i}$$

- **Función de autocovarianza:**

$$\gamma(h) = \sum_{i=1}^{p} \alpha_i \gamma(h-i) + \sum_{i=h+1}^{p} \mu(h-i, \alpha_i) + \delta_h(0)\sigma_\epsilon^2$$

para $h \geq 0$, donde $\delta_h(0) = 1$ si $h = 0$ y 0 en otro caso. Además, se tiene que:

$$\mu(h-i, \alpha_i) = \sum_{j=1}^{i+1-h} \alpha_j \mu_{j+h-i, \alpha_i} + \alpha_i(\delta_i(i-h) - \alpha_{i-h})\mu_x$$

Si $h < i$ y 0 en otro caso, siendo $\mu_x = E[X_t]$.

Propiedades

Du y Li (1991) introducen una variante del proceso INAR(P) al considerar que las componentes del vector $(\alpha_1 \circ X_t, \alpha_2 \circ X_t, \cdots, \alpha_p \circ X_t | x_t)$ son independientes. Esto hace que su definición sea análoga a la del proceso AR(P) y que la estructura de autocovarianza esté dada por:

$$\gamma(h) = \alpha_1\gamma(h-1) + \alpha_2\gamma(h-2) + \cdots + \alpha_p\gamma(h-p)$$
$$\rho(h) = \alpha_1\rho(h-1) + \alpha_2\rho(h-2) + \ldots + \alpha_p\rho(h-p)$$

Estimaciones

Estimadores de Yule-Walker

Este método consiste en reemplazar la función de autocovarianza teórica por la función de autocovarianza muestral, es decir.

$$\begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \dots & \hat{\gamma}(p-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \dots & \hat{\gamma}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(p-1) & \hat{\gamma}(p-2) & \dots & \hat{\gamma}(0) \end{bmatrix} \begin{bmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \vdots \\ \widehat{\alpha}_p \end{bmatrix} = \begin{bmatrix} \hat{\gamma}(1) \\ \hat{\gamma}(2) \\ \vdots \\ \hat{\gamma}(p) \end{bmatrix}$$

Por lo que los estimadores para $\alpha_1, \cdots, \alpha_p$, se obtienen al resolver el sistema de ecuaciones lineales en el orden de los parámetros del sistema anterior.

Estimaciones

Para obtener el estimador de λ, nótese que:

$$\epsilon_t = X_t - \alpha_1 \circ X_{t-1} - \cdots - \alpha_p \circ X_{t-p}$$
$$E[\epsilon_t] = E[X_t] - \alpha_1 E[X_{t-1}] - \cdots - \alpha_p E[X_{t-p}]$$

Un estimador razonable para λ es:

$$\widehat{E[\epsilon_t]} = \frac{1}{n}\left(\sum_{t=1}^{n} X_t - \widehat{\alpha_1}\sum_{t=1}^{n} X_{t-1} - \cdots - \widehat{\alpha_p}\sum_{t=1}^{n} X_{t-p}\right)$$
$$= \frac{1}{n}\sum_{t=1}^{n}(X_t - \widehat{\alpha_1}X_{t-1} - \cdots - \widehat{\alpha_p}X_{t-p})$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Estimaciones

Estimadores Mínimos Cuadrados Condicionales

El procedimiento de estimación se basa en minimizar la suma de las desviaciones al cuadrado respecto a la media condicional. Así, los estimadores mínimos cuadrados condicionales para θ son los valores que minimizan:

$$\sum_{t=p}^{n} (X_t - E\left[X_t / \left(X_{t-1}, \cdots, X_{t-p}\right)\right])^2$$

Para $\theta = (\alpha_1, \alpha_2, \ldots, \alpha_p, \lambda)^T$ conjunto de parámetros a estimar. En este orden de ideas, diferenciando con respecto a θ_i para $i = 1, 2, \ldots, (p+1)$ e igualando a cero, se obtienen los estimadores de mínimos cuadrados condicionales $\hat{\theta}_i$.

UNIVERSIDAD NACIONAL DE COLOMBIA

INMA(1)

El modelo de Medias Móviles de valores enteros no negativos de primer orden, **INMA(1)**. Sea $\{X_t;\ t = \pm 1, \pm 2, ...\}$ una sucesión de v.a. enteras no-negativas, este proceso se define como

$$X_t = \beta \circ \epsilon_{t-1} + \epsilon_t$$

Donde $\beta \in [0, 1]$ y ϵ_t, $t \in \mathbb{Z}$ es una sucesión de v. a. no-negativas (i.i.d.) con media $\mu_\epsilon < \infty$ y varianza $\sigma^2_\epsilon < \infty$.

UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades

Tal como los MA(1), los procesos INMA(1) son procesos con memoria de un único periodo, es decir, cualquier valor de X_t esta corrrelacionado con X_{t-1} y X_{t+1} y con ningún otro valor de la serie.

Propiedades

Como las v. a. $\beta \circ \epsilon_{t-i}$ y $\beta \circ \epsilon_{t-j}$ son independientes para $i \neq j$, el proceso INMA(1) es estacionario porque:

- $E[X_t] = E[\beta \circ \epsilon_{t-1} + \epsilon_t] = (1+\beta)\mu_\epsilon$
- $V[\beta \circ \epsilon_{t-1}] = \beta(1-\beta)\mu_\epsilon + (1-\beta^2)\sigma_\epsilon^2$
- $\gamma(h) = \begin{cases} \beta\sigma_\epsilon^2 & \text{si} \quad h = 1 \\ 0 & \text{si} \quad h > 1 \end{cases}$
- $\rho(h) = \begin{cases} \frac{\beta\sigma_\epsilon^2}{\beta(1-\beta)\mu_\epsilon + (1+\beta^2)\sigma_\epsilon^2} & \text{si} \quad h = 1 \\ 0 & \text{si} \quad h > 1 \end{cases}$

Estimaciones

Estimadores por mínimos cuadrados condicionales

Tenemos el modelo INMA(1) el cual está dado por:

$$X_t = \beta \circ \epsilon_{t-1} + \epsilon_t$$

para obtener los estimadores de β y λ de un modelo INMA(1), hacemos uso de la expresión:

$$E[X_t|\epsilon_{t-1}] = \beta\epsilon_{t-1} + \lambda$$

los estimadores de mínimos cuadrados condicionales (CLS) para β y λ son los valores que minimizan:

$$Q_n(\beta, \lambda) = \sum_{t=1}^{n}(X_t - E[X_t|\epsilon_{t-1}])^2$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Estimaciones

Estimadores por mínimos cuadrados condicionales

Asumiendo una marginal Poisson

$$Q_n(\beta,\lambda) = \sum_{t=1}^{n}(X_t - E[X_t|\epsilon_{t-1}])^2 = \sum_{t=1}^{n}(X_t - [\beta\epsilon_{t-1}+\lambda])^2$$

$$= \sum_{t=1}^{n}(X_t^2 - 2X_t(\beta\epsilon_{t-1}+\lambda) + (\beta\epsilon_{t-1}+\lambda)^2)$$

$$= \sum_{t=1}^{n}(X_t^2 - 2\beta X_t\epsilon_{t-1} - 2\lambda X_t + \beta^2\epsilon_{t-1}^2 + 2\beta\lambda\epsilon_{t-1} + \lambda^2)$$

Estimaciones

Estimadores por mínimos cuadrados condicionales

derivando $Q_n(\beta, \lambda)$ con respecto a β y λ tenemos:

$$\frac{\partial Q_n(\beta, \lambda)}{\partial \beta} = \sum_{t=1}^{n} (-2X_t \epsilon_{t-1} + 2\beta \epsilon_{t-1}^2 + 2\lambda \epsilon_{t-1}) = 0$$

$$\frac{\partial Q_n(\beta, \lambda)}{\partial \lambda} = \sum_{t=1}^{n} (-2X_t + 2\beta \epsilon_{t-1} + 2\lambda) = 0$$

luego

$$\hat{\lambda} = \frac{1}{n} \left[\sum_{t=1}^{n} X_t - \hat{\beta} \sum_{t=1}^{n} \epsilon_{t-1} \right]$$

$$\hat{\beta} = \frac{n \sum_{t=1}^{n} X_t \epsilon_{t-1} - \sum_{t=1}^{n} X_t \sum_{t=1}^{n} \epsilon_{t-1}}{n \sum_{t=1}^{n} \epsilon_{t-1}^2 - (\sum_{t=1}^{n} \epsilon_{t-1})^2}$$

INMA(q)

Al-Osh y Alzaid (1988) hacen una geralización para el proceso INMA(1), considerando un proceso en el que la máxima vida para un elemento en el sistema es $(q+1)$ unidades de tiempo.
Se consideran, α_j con $j = 1, 2, ..., q$ como la probabilidad que tiene un elemento de ϵ_t de aparecer por primera vez en el tiempo $t + j$ y ser un elemento de X_{t+j}, donde $\sum_{j=1}^{q} \alpha_j \leq 1$.

Se definen los coeficientes β_i con $i = 0, 1, 2, ..., q$ como:

$$\beta_0 = 1$$
$$\beta_i = \sum_{k=1}^{i} \alpha_k \beta_{i-k} \text{ para i} = 1, 2, ..., q$$

β_i es la probabilidad que tiene un elemento de ϵ_t de ser un elemento de X_{t+i}.

INMA(q)

Por otra parte, sea $\{\mathbb{Y}_i^t = (Y_{i,1}^t, Y_{i,2}^t, ..., Y_{i,q}^t); i = 1, 2, ..., t \in \mathbb{Z}\}$ una sucesión de vectores i.i.d. donde:

$$Y_{i,j}^t = \begin{cases} 1 & \text{si el i-ésimo elemento de } \epsilon_t \\ & \text{está presente en el sistema en el tiempo } t+j \\ 0 & \text{En otro caso.} \end{cases}$$

entonces $\beta_j = P(Y_{i,j}^t = 1)$ para $j = 1, 2, ..., q$ y para todo $i = 1, 2, ..., y$ y $t = 0, \pm 1, \pm 2, ...$ Luego tenemos que:

$$P(Y_{i,\kappa_1}^t = 1, Y_{i,\kappa_2}^t = 1, ..., Y_{i,\kappa_j}^t = 1) = \prod_{l=1}^{j} \beta_{\kappa_l - \kappa_{l-1}}$$

Con $\kappa_0 = 0$

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelo INMA(q)

Teniendo en cuenta las definiciones anteriores, el proceso INMA(q) se define:

$$X_t = \beta_1 \circ \epsilon_{t-1} + \beta_2 \circ \epsilon_{t-2} + ... + \beta_q \circ \epsilon_{t-q} + \epsilon_t$$

donde $\{\epsilon_t, t \in \mathbb{Z}\}$ es una sucesión de v.a. (i.i.d.) con media μ_ϵ y varianza σ^2_ϵ finita y

$$\beta_j \circ \epsilon_t = \sum_{i=1}^{\epsilon_t} Y^t_{i,j}$$

que denota los sobrevivientes de ϵ_t en el tiempo $t + j$.

UNIVERSIDAD NACIONAL DE COLOMBIA

Propiedades

Suponiendo que este modelo es siempre estacionario y que ϵ_t tiene media μ_ϵ y varianza finita σ_ϵ^2

- $E(X_t) = \mu_\epsilon \sum_{j=0}^{q} \beta_j$
- $V(X_t) = \mu_\epsilon \sum_{j=0}^{q} \beta_j(1-\beta_j) + \sigma_\epsilon^2 \sum_{j=0}^{q} \beta_j^2$
- $$\gamma(h) = \begin{cases} \sum_{j=0}^{q-h} \left[\beta_j(\beta_h - \beta_{h+j})\mu_\epsilon + \beta_j \beta_{h+j}\sigma_\epsilon^2\right] & \text{si} \quad h = 1, 2, ..., q \\ 0 & \text{si} \quad h > q \end{cases}$$
- $$\rho(h) = \begin{cases} 1 & \text{si} \quad h = 0 \\ \frac{\sum_{j=0}^{q-h}\left[\beta_j\left(\beta_h - \beta_{j+h}\right)\mu_\epsilon + \beta_j\beta_{j+h}\sigma_\epsilon^2\right]}{\mu_\epsilon \sum_{j=0}^{q} \beta_j\left(1-\beta_j\right) + \sigma_\epsilon^2 \sum_{j=0}^{q}\beta_j^2} & \text{si} \quad 1 \le h \le q. \\ 0 & \text{si} \quad h > q \end{cases}$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Estimaciones

Estimadores por mínimos cuadrados condicionales

Tenemos el modelo INMA(q) el cual está dado por:

$$X_t = \beta_1 \circ \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + ... + \beta_q \circ \epsilon_{t-q} + \epsilon_t$$

tenemos

$$E[X_t|(\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_{t-p})] = \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + ... + \beta_p \epsilon_{t-p} + \lambda$$

los estimadores mínimos cuadrados condicionales para θ son los valores que minimizan:

$$\sum_{t=p}^{n} (X_t - E[X_t|(\epsilon_{t-1}, \epsilon_{t-2}, ..., \epsilon_{t-p})])^2$$

Modelo INARMA

Un proceso INARMA (Promedio móvil Autorregresivo para valores enteros) es una extensión del modelo ARMA aplicado a series de tiempo donde las observaciones son valores enteros no negativos.

Este proceso mixto se puede ver cómo la unión de los procesos INAR e INMA, siempre que tengan un proceso de innovación común $\{\epsilon_t, t \in \mathbb{Z}\}$.

Modelo INARMA(1,1)

Este proceso se construye uniendo dos modelos, el INAR(1) y el INMA(1), con un proceso de innovación en común, el cual es una sucesión de variables aleatorias *i.i.d* con media μ_ϵ y varianza σ_ϵ^2. El proceso se define como:

$$X_t = \alpha \circ X_{t-1} + \beta \circ \epsilon_{t-1} + \epsilon_t$$

donde β se define como en el proceso INMA(1). Además, sí $\alpha \in [0,1), \mu_\epsilon < \infty, \sigma_\epsilon^2 < \infty$, entonces X_t es estacionario.

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelo INARMA(1,1)

Propiedades

Algunas de las propiedades básicas de este modelo son

- Valor esperado

$$E[X_t] = \frac{(1+\beta)\mu_\epsilon}{1-\alpha}$$

- Varianza

$$V(X_t) = \frac{(\alpha(1+\beta)+\beta(1-\beta)\mu_\epsilon)+(\beta(\beta+2\alpha)+1)\sigma_\epsilon^2}{1-\alpha^2}$$

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelo INARMA(1,1)

Propiedades

- Función de autocovarianza

$$\gamma(h) = \begin{cases} V(X_t) & \text{si h=0} \\ \alpha^h V(X_t) + \alpha^{h-1}\beta\sigma_\epsilon^2 & \text{si h} \geq 1 \end{cases}$$

- Función de autocorrelación la obtenemos de dividir $\gamma(h)$ entre $V(X_t)$.

UNIVERSIDAD NACIONAL DE COLOMBIA

Modelo INARMA(1,q)

Este modelo se construye uniendo dos modelos, el INAR(1) y el INMA(q) con un proceso de innovación en común $\{\epsilon_t, t \in \mathbb{Z}\}$, el cual es una sucesión de variables aleatorias *i.i.d* con media μ_ϵ y varianza σ^2_ϵ. El proceso se define como:

$$X_t = \alpha \circ X_{t-q} + \sum_{i=1}^{q} \beta_i \circ \epsilon_{t+1-i} + \epsilon_t$$

Los β_i se definen como en el proceso INMA(q). Además, si $\alpha \in [0, 1)$, $\mu_\epsilon < \infty$, $\sigma^2_\epsilon < \infty$, entonces el proceso X_t es estacionario.

Modelo INARMA(1,q)

Propiedades

- Valor esperado

$$E[X_t] = \frac{\mu_\epsilon}{1-\alpha} + \mu_\epsilon \sum_{i=1}^{q} \beta_i$$

- Varianza

$$V(X_t) = \frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2} + \sum_{i=1}^{q} [\beta_i(1-\beta_i)\mu_\epsilon + \beta_i^2\sigma_\epsilon^2]$$

Modelo INARMA(1,q)

Propiedades

- Función de autocovarianza

$$\gamma(h) = \begin{cases} \left(\frac{\alpha\mu_\epsilon+\sigma_\epsilon^2}{1-\alpha^2}\right)\alpha^h + \sum_{i=1}^{q-h}(\beta_i(\beta_h-\beta_{h+i})\mu_\epsilon + \beta_i\beta_{h+i}\sigma_\epsilon^2) & \\ +\sigma_\epsilon^2 \sum_{i=q-h+1}^{q} \alpha^{h-1+i-q}\beta_i & \text{si } 1 \leq h \leq q \\ \left(\left(\frac{\alpha\mu_\epsilon+\sigma_\epsilon^2}{1-\alpha^2}\right)\alpha^q + \sigma_\epsilon^2\sum_{i=1}^{q}\alpha^{i-1}\beta_i\right)\alpha^{h-q} & \text{si } h > q \end{cases}$$

- La función de autocorrelación la obtenemos de dividir $\gamma(h)$ entre $V(X_t)$.

Modelo INARMA(1,q) con marginal de Poisson

Adicional a lo que se definió para el modelo INARMA(1,q), se tiene que el supuesto de que $\epsilon_t \sim$ *Pois*(λ) para todo t, entonces:

- X_t también se distribuye Poission con media $\mu = \lambda(\frac{1}{1-\alpha} + \sum_{i=1}^{q} \beta_i)$.

Si se tiene el caso particular en que $q = 1$, es decir, si se tiene un modelo INARMA(1,1), entonces la media de la distribución Poisson será $\mu = -\frac{\lambda(1+(1-\alpha)\beta)}{1-\alpha}$.

Series de tiempo categóricas

Las series de tiempo categóricas pueden contener rangos cualitativos de un número finito de categorías y en ocasiones ese rango puede considerarse ordinal o nominal.

Modelo de Transición de Estado

En este tipo de modelo, se asume que la variable de interés puede estar en un conjunto finito de estados, y que la probabilidad de pasar de un estado a otro depende solo del estado actual y no de los estados anteriores.Un modelo de transición de estado se puede representar mediante una matriz de transición de estado, que especifica las probabilidades de pasar de un estado a otro en un solo paso de tiempo.

Modelo de Markov parsimoniosamente parametrizado

Es un modelo de transición de estado usado para reducir la cantidad de parámetros necesarios en un modelo de Markov.
Se parametriza por medio de la matriz de transición de estado.

Modelos lineales Generalizados

- Estos modelos son análogos a los modelos de heteroscedasticidad condicional autorregresiva generalizada (GARCH).
- Algunas ventajas que tienen los modelos basados en GLM con respecto a los basados en el operador de refinamiento son:
  1. Estos pueden describir efectos de covariables y correlaciones negativas de forma sencilla.
  2. Hay un rico conjunto de herramientas para esta clase de modelos.

Modelos lineales Generalizados (modelo)

Se está interesado en modelos de la forma:

$$g(\lambda_t) = \beta_0 + \sum_{k=1}^{p} \beta_k \tilde{g}(Y_{t-i_k}) + \sum_{l=1}^{q} \alpha_l g(\lambda_{t-i_l}) + \eta^T X_t$$

Donde se tiene que:

- $g(\lambda_t)$:predictor lineal.
- $\{Y_t : t \in \mathbb{N}\}$:Serie de tiempo.
- X_t: vector de covariables.
- $E[Y_t|\mathcal{F}_{t-1}] = \lambda_t$.
- $\mathcal{F}_t$: La historia del proceso conjunto $\{Y_t, \lambda_t, X_{t+1} : t \in \mathbb{N}\}$.
- $g : \mathbb{R}^+ \to \mathbb{R}$: función de enlace.
- $\tilde{g} : \mathbb{N}_0 \to \mathbb{R}$: función de transformación.

Modelos lineales Generalizados (modelo)

- Consideremos $g(x) = \tilde{g}(x) = x$ el modelo se transforma en:

$$\lambda_t = \beta_0 + \sum_{k=1}^{p} \beta_k Y_{t-k} + \sum_{l=1}^{q} \alpha_l \lambda_{t-l}$$

- Consideremos $g(x) = Log(x)$ y $\tilde{g}(x) = log(x+1)$ el modelo se transforma en (modelo log lineal):

$$\nu_t = \beta_0 + \sum_{k=1}^{p} \beta_k log(Y_{t-k} + 1) + \sum_{l=1}^{q} \alpha_l \nu_{t-l}$$

donde $\nu_t = log(\lambda_t)$

Modelos lineales Generalizados (modelo)

- Dependiendo la distribución que asumamos para $Y_t|\mathcal{F}_{t-1}$, obtendremos distintas expresiones para la media y la varianza condicionales.
- se hace la estimación vía máxima verosimilitud cuasi condicional, para un vector de observaciones $y = (y_1, \cdots, y_n)^T$, la función de cuasi log-verosimilitud se puede escribir como:

$$l(\theta) = \sum_{t=1}^{n} P_t(y_t; \theta) = \sum_{t=1}^{n} (y_t ln(\lambda_t(\theta)) - \lambda_t(\theta))$$

El estimador de quasi maxima-verosimilitud $\hat{\theta}$ de θ, asumiendo que existe es la solución del problema de optimización $\hat{\theta} = \hat{\theta}_n = argmax_{\theta \in \Theta} l(\theta)$

Referencias

- Funes Torres, J. N. (2010). Modelos INARMA para series temporales de valores enteros: análisis, propiedades asintóticas y estimación. Ene, 12, 16.
- Weiß, C. H. (2018). An introduction to discrete-valued time series. John Wiley Sons.
- Gonzalez Lopez, C. F., (2011) Modelos de Series de Tiempo para Datos Enteros.
- Argueta Abarca, J. E. (s. f.). Modelos para series temporales de valores enteros [Tesis de grado]. Universidad de El Salvador

Referencias

- Bracher, J. (2019). A New INARMA(1, 1) Model with Poisson Marginals. In: Steland, A., Rafajłowicz, E., Okhrin, O. (eds) Stochastic Models, Statistics and Their Applications. SMSA 2019. Springer Proceedings in Mathematics Statistics, vol 294. Springer, Cham.
- MacDonald, I. L. (s. f.). Time Series Models for Discrete Data [Tesis doctoral]. University of Cape Town
- Liboschik, T., Fokianos, K., Fried, R. (2017). Tscount: An R package for analysis of count time series following generalized linear models. Journal of statistical software, 82(5). https://doi.org/10.18637/jss.v082.i05

UNIVERSIDAD NACIONAL DE COLOMBIA

Muchas gracias por la atención :3