



UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA



Universidad de El Salvador
Hacia la libertad por la cultura

TRABAJO DE GRADUACIÓN TITULADO:
“MODELOS PARA SERIES TEMPORALES DE VALORES ENTEROS”

TRABAJO DE GRADUACIÓN PRESENTADO POR:
Br. JOSÉ ELÍAS ARGUETA ABARCA.

PARA OPTAR AL GRADO DE:
LICENCIADO EN ESTADÍSTICA.

CIUDAD UNIVERSITARIA, JUNIO DE 2010.



UNIVERSIDAD DE EL SALVADOR

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

ESCUELA DE MATEMÁTICA



Universidad de El Salvador
Hacia la libertad por la cultura

PROYECTO DE GRADUACIÓN TITULADO:
“MODELOS PARA SERIES TEMPORALES DE VALORES ENTEROS”

TRABAJO DE GRADUACIÓN PRESENTADO POR:
Br. JOSÉ ELÍAS ARGUETA ABARCA.

PARA OPTAR AL GRADO DE:
LICENCIADO EN ESTADÍSTICA.

ASESOR: _____
Dr. JOSÉ NERYS FUNES TORRES.

CIUDAD UNIVERSITARIA, JUNIO DE 2010.

UNIVERSIDAD DE EL SALVADOR

RECTOR:

Ing. RUFINO ANTONIO QUEZADA SÁNCHEZ.

SECRETARIO GENERAL:

Lic. DOUGLAS VLADIMIR ALFARO CHÁVEZ.

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO:

Dr. RAFAEL ANTONIO GÓMEZ ESCOTO.

SECRETARIA:

Licda. MARIA TRINIDAD TRIGUEROS DE CASTRO.

ESCUELA DE MATEMÁTICA

DIRECTOR:

Ing. CARLOS MAURICIO CANJURA.

A Dios, mi gran familia y amigos.

Índice general

1. Propiedades Generales de los Modelos para Series Temporales de Valores Enteros.	1
1.1. Introducción.	1
1.2. Operador de Refinamiento.	4
1.3. Modelo Autorregresivo de Valores Enteros.	8
1.3.1. Definición del Modelo Autorregresivo de orden 1, INAR(1).	8
1.3.2. Propiedades Generales del Modelo Autorregresivo de orden 1, INAR(1).	9
1.3.3. Definición del Modelo Autorregresivo de orden 2, INAR(2).	13
1.3.4. Propiedades Generales del Modelo Autorregresivo de orden 2, INAR(2).	13
1.3.5. Definición del Modelo Autorregresivo de orden p , INAR(p).	15
1.3.6. Propiedades Generales del Modelo Autorregresivo de orden p , INAR(p).	16
1.3.7. Estacionariedad.	18
1.4. Modelo de Media Móvil de Valores Enteros.	21
1.4.1. Definición del Modelo de Media Móvil de orden 1, INMA(1).	21
1.4.2. Propiedades Generales del Modelo de Media Móvil de orden 1, INMA(1).	21
1.4.3. Función Generatriz de Probabilidad del Modelo INMA(1).	24
1.4.4. Aspectos Básicos de un Proceso INMA(1) de Poisson.	28
1.4.5. Invertibilidad.	30
1.4.6. Definición del Modelo de Media Móvil de orden 2, INMA(2).	33
1.4.7. Propiedades Generales del Modelo de Media Móvil de orden 2, INMA(2).	34
1.4.8. Definición del Modelo de Media Móvil de orden q , INMA(q).	38
1.4.9. Propiedades Generales del Modelo de Media Móvil de orden q , INMA(q).	39

1.5. Modelo mixto Autorregresivo y de Media Móvil de Valores Enteros de orden	
1. INARMA(1,1).	43
1.5.1. Definición del Modelo mixto Autorregresivo y de Media Móvil de Va-	
lores Enteros de orden 1. INARMA(1,1).	43
1.5.2. Propiedades Generales del Modelo INARMA(1,1).	44
2. Estimación de Parámetros de los Modelos INARMA.	49
2.1. Estimación de Parámetros del Modelo INAR(1) con marginal de Poisson. . .	49
2.1.1. Estimadores de Yule-Walker.	49
2.1.2. Estimadores Mínimos Cuadrados Condicionales.	51
2.1.3. Estimadores de Máxima Verosimilitud.	53
2.2. Estimación de Parámetros del Modelo INAR(2) con marginal de Poisson. . .	60
2.2.1. Estimadores de Yule-Walker.	60
2.2.2. Estimadores Mínimos Cuadrados Condicionales.	62
2.2.3. Estimadores de Máxima Verosimilitud.	68
2.3. Estimación de Parámetros del Modelo INAR(p) con marginal de Poisson. . .	75
2.3.1. Estimadores de Yule-Walker.	75
2.3.2. Estimadores Mínimos Cuadrados Condicionales.	77
2.4. Estimación de Parámetros del Modelo INMA(1) con marginal de Poisson. . .	78
2.4.1. Estimadores Mínimos Cuadrados Condicionales.	78
2.5. Estimación de Parámetros del Modelo INMA(2) con marginal de Poisson. . .	80
2.5.1. Estimadores Mínimos Cuadrados Condicionales.	80
2.6. Estimación de Parámetros del Modelo INMA(q) con marginal de Poisson. . .	86
2.6.1. Estimadores Mínimos Cuadrados Condicionales.	86
3. Estudio de simulación y aplicaciones reales.	87
3.1. Simulación del modelo INAR	87
3.2. Simulación del modelo INMA	93
3.3. Aplicación de datos reales.	96
3.3.1. Casos de Dengue.	96

ÍNDICE GENERAL _____ III

3.3.2. Número de Partos. 98

Conclusiones **103**

Bibliografía **105**

Índice de cuadros

3.1. INAR(1). Yule-Walker	88
3.2. INAR(1). Mínimos Cuadrados Condicionales	89
3.3. INAR(1). Máxima Verosimilitud	90
3.4. INAR(2). Yule-Walker	91
3.5. INAR(2). Mínimos Cuadrados Condicionales	92
3.6. INMA(1). Mínimos Cuadrados Condicionales	93
3.7. INMA(2). Mínimos Cuadrados Condicionales	94
3.8. Estimación de parámetros del modelo INAR(5), ajustado al modelo (3.2) . .	98

Índice de figuras

3.1. Número de casos diarios de Dengue durante el 2005. <i>Fuente:</i> Hospital I.S.S.S. de Amatepec, Soyapango.	96
3.2. Función de autocorrelación para la serie temporal Dengue.	96
3.3. Función de autocorrelación para los errores del modelo (3.1).	97
3.4. Histograma de los errores de la ecuación (3.1).	97
3.5. Número de partos diarios durante el 2006. <i>Fuente:</i> Hospital I.S.S.S. Amatepec, Soyapango.	98
3.6. Función de autocorrelación para la serie temporal Partos.	99
3.7. Histograma de los errores de la ecuación (3.3).	99
3.8. Función de autocorrelación de los errores de la ecuación (3.3).	100
3.9. Histograma de los errores de la ecuación (3.4).	100
3.10. Función de Autocorrelación de los errores de la ecuación (3.4).	101

Capítulo 1

Propiedades Generales de los Modelos para Series Temporales de Valores Enteros.

1.1. Introducción.

Las series temporales aparecen en muchas situaciones como por ejemplo, la compra o venta mensual de un producto, la temperatura diaria en una región, el consumo mensual de energía eléctrica, tasas anuales de crecimiento de la población, etc. estos modelos de tipo ARMA presentan un comportamiento continuo en los datos. Pero en muchas de las situaciones, las observaciones registradas son valores enteros como por ejemplo, el número de accidentes ocurridos diariamente en un lugar específico, el número de pacientes que ingresan a un hospital por cierta enfermedad, el número de unidades defectuosas registradas diariamente en un proceso de producción, etc. en estos casos surge la necesidad de utilizar modelos que reflejen la naturaleza de dichos datos.

Antes de definir los modelos de series temporales de valores enteros, se define el operador de refinamiento y sus propiedades, definido por Steutel y Van Harn (1979), la idea consiste en reemplazar la multiplicación escalar de los modelos ARMA estándar por dicho operador “*”; luego, se estudian las propiedades generales de los modelos autorregresivos, es decir, la media

o valor esperado, la varianza, la covarianza y la correlación, éstas propiedades discutidas por Al-Osh y Alzaid (1987), Alzaid y Al-Osh (1990), Du y Li (1991), Gauthier y Latour (1994), Park y Oh (1997), Latour (1998), Funes (2001), Marques da Silva (2005); de igual manera, se estudian las mismas propiedades para los modelos de media móvil, estudiados por Alzaid y Al-Osh (1988b), McKenzie (1988b), Park y Kim (1997), Funes (2001); por último, se estudian las propiedades generales del modelo mixto autorregresivo y de media móvil de orden uno, estudiado por McKenzie (1988b) para el proceso de Poisson ARMA(1, q) y por Funes (2001).

Superada la fase de definir las propiedades generales de un modelo INARMA, la siguiente etapa es estimar los parámetros del proceso INAR(1), INAR(p), INMA(1) e INMA(q), por lo que se presentan los resultados obtenidos por Al-Osh y Alzaid (1987), ellos obtienen los estimadores de los parámetros del tipo Yule-Walker, Mínimos Cuadrados Condicionales y Máxima Verosimilitud Condicional para el proceso INAR(1) con marginal de Poisson, Marques da Silva (2005) obtiene los estimadores de Yule-Walker del proceso INAR(p) definido por Al-Osh y Alzaid (1987), Du y Li (1991) estudian el proceso INAR(p) y obtienen los estimadores de los parámetros usando el método de Mínimos Cuadrados Condicionales, Funes (2001) obtiene los estimadores del proceso INAR(p) definido por Al-Osh y Alzaid (1987), e INMA(q) definido por McCormick y Park (1997); por otra parte, YouSung Park y MyungJin Kim (1997) obtienen los estimadores del proceso INMA(q) con marginal de Poisson, éstos últimos observan que los estimadores de los parámetros λ y β_i para $i = 1, 2, \dots, q$ pueden obtenerse a través de la función de autocovarianza estimada.

Para evaluar la fiabilidad de los métodos de estimación de parámetros descritos, se presentan unas simulaciones de los modelos INAR(1), INAR(2), INMA(1) e INMA(2), donde suponemos para cada modelo, que ϵ_t tiene una distribución de Poisson de parámetro λ , para realizar el experimento se consideran modelos de tamaño, $n = 50, 75, 100, 150$ y 200 , con parámetros α_i y β_i entre 0.1 y 0.9, con $i = 1, 2$ (según el tipo y orden del modelo), para cada combinación de parámetros se realizan 100 replicaciones; se calculan el sesgo (que se define como *una medida que cuantifica la distancia entre el valor esperado del estimador y el verdadero valor del parámetro*) y el error cuadrático medio (que se define como *una medida que cuantifica la concentración de las estimaciones alrededor del verdadero valor del parámetro*)

para cada combinación. Las simulaciones se realizaron usando el software R-project.

Por otra parte, se presenta el ajuste de los modelos a dos series de datos:

1. El número de pacientes atendidos diariamente en el Hospital I.S.S.S. (Instituto Salvadoreño del Seguro Social) de Amatepec, El Salvador y confirmados con El Dengue, entre el 11 de agosto y el 17 de diciembre de 2005 (ver figura 3.1).
2. El número de partos atendidos diariamente (lunes a viernes) en el Hospital I.S.S.S. (Instituto Salvadoreño del Seguro Social) de Amatepec, El Salvador, en el año 2006 (ver figura 3.5).

1.2. Operador de Refinamiento.

La definición del operador ‘*’ fue introducida por Steutel y Van Harn (1979) y también por Van Harn (1978, P. 85).

DEFINICIÓN 1.1 Sea X una variable aleatoria v.a. entera y no-negativa; entonces para todo $\alpha \in [0, 1)$ el operador ‘*’ está definido por:

$$\alpha * X = \sum_{i=1}^X Y_i$$

Donde Y_i , representa una serie de conteos y es una sucesión de v.a. independientes e idénticamente distribuidas i.i.d. e independiente de X , tal que $E(Y_i) = \alpha$ y $V(Y_i) = \sigma^2$. Esta es una generalización del operador de refinamiento binomial, donde los conteos de la serie son una secuencia de v.a. i.i.d. de Bernulli, tal que:

$$P(Y_i = 1) = 1 - P(Y_i = 0) = \alpha$$

Es obvio que, dado X , $\alpha * X$ tiene una distribución binomial con parámetro $Bin(\alpha, X)$.

Del operador de refinamiento, se tienen las siguientes propiedades:

1. $0 * X = 0$

Sea $0 * X \Rightarrow \alpha = 0$, es decir $P(Y_i = 1) = 1 - P(Y_i = 0) = 0$, entonces $P(Y_i = 0) = 1$

2. $1 * X = X$

Sea $1 * X \Rightarrow \alpha = 1$, es decir $P(Y_i = 1) = 1 - P(Y_i = 0) = 1$, entonces $P(Y_i = 0) = 0$

3. $E(\alpha * X) = \alpha E(X)$

En efecto:

Aplicando la propiedad de valor esperado a la definición (1.1), tenemos:

$$E(\alpha * X) = E\left(\sum_{i=1}^X Y_i\right)$$

dado que X es una v.a. no sabemos hasta donde puede tomar valores la sumatoria, pero sabemos que si condicionamos la sumatoria por X suponiendo conocido ese valor, tenemos:

$$E\left(\sum_{i=1}^X Y_i\right) = E\left(E\left(\sum_{i=1}^X Y_i / X\right)\right)$$

dato que Y_i tiene distribución Bernoulli con parámetros α , entonces $\sum_{i=1}^X Y_i/X$, tiene una distribución binomial con parámetros α y X :

$$E\left(E\left(\sum_{i=1}^X Y_i/X\right)\right) = E(\alpha X) = \alpha E(X) \quad \blacksquare$$

4. $E((\alpha * X)X) = \alpha E(X^2)$

En efecto:

$$\begin{aligned} E((\alpha * X)X) &= E(E((\alpha * X)X/X)) = E(XE((\alpha * X)/X)) = E(X\alpha X) \\ &= \alpha E(X^2) \quad \blacksquare \end{aligned}$$

5. Sea $\beta \in [0, 1]$, entonces $E((\alpha * X)(\beta * X)) = \alpha\beta E(X^2)$

En efecto:

$$E((\alpha * X)(\beta * X)) = E(E((\alpha * X)(\beta * X)/X))$$

supondremos que $(\alpha * X)$ es independiente de $(\beta * X)$, dado el valor de X .

$$\begin{aligned} E(E((\alpha * X)(\beta * X)/X)) &= E(E((\alpha * X)/X)E((\beta * X)/X)) \\ &= E((\alpha X)(\beta X)) = \alpha\beta E(X^2) \quad \blacksquare \end{aligned}$$

6. Para cualquier $\beta \in [0, 1]$, se verifica que:

$$\beta * (\alpha * X) \stackrel{d}{=} (\beta\alpha) * X$$

Prueba:

Utilizando la función generatriz de probabilidad, la cual se define por:

$$\phi_X(s) = E(s^X), \text{ donde } |s| < 1:$$

a) Sea $w_1 = \beta * (\alpha * X)$

La función generatriz de probabilidad de w_1 es:

$$\phi_{w_1}(s) = E(s^{\beta * (\alpha * X)})$$

por la definición (1.1) del operador, tenemos:

$$\phi_{w_1}(s) = E \left(s^{\beta * \left(\sum_{i=1}^X Y_i \right)} \right) = E \left(E \left(s^{\beta * \left(\sum_{i=1}^X Y_i / X \right)} \right) \right)$$

si hacemos $Z = \left(\sum_{i=1}^X Y_i / X \right) \sim Bin(\alpha, X)$, tenemos:

$$\phi_{w_1}(s) = E \left(E \left(s^{\beta * \left(\sum_{i=1}^X Y_i / X \right)} \right) \right) = E \left(E \left(s^{\beta * (Z)} \right) \right)$$

nuevamente, si aplicamos la definición (1.1) del operador, tenemos:

$$\phi_{w_1}(s) = E \left(E \left(s^{\beta * (Z)} \right) \right) = E \left(E \left(E \left(s^{\left(\sum_{i=1}^Z Y_i / Z \right)} \right) \right) \right)$$

si llamamos $Z' = \left(\sum_{i=1}^Z Y_i / Z \right) \sim Bin(\beta, Z)$, tenemos:

$$\begin{aligned} \phi_{w_1}(s) &= E \left(E \left(E \left(s^{\left(\sum_{i=1}^Z Y_i / Z \right)} \right) \right) \right) = E \left(E \left(E \left(s^{(Z')} \right) \right) \right) \\ &= E \left(E \left((1 - \beta)s^0 + \beta s^1 \right)^Z \right) = E \left(E \left(1 - \beta + \beta s \right)^Z \right) \\ &= E \left((1 - \alpha) (1 - \beta + \beta s)^0 + \alpha (1 - \beta + \beta s)^1 \right) \\ &= E \left((1 - \alpha) + \alpha (1 - \beta + \beta s)^1 \right) = E \left((1 - \alpha) + \alpha - \alpha\beta + \alpha\beta s \right) \\ &= E \left(1 - \alpha\beta + \alpha\beta s \right) \quad \blacksquare \end{aligned}$$

b) Sea $w_2 = (\beta\alpha) * X$

La función generatriz de probabilidad de w_2 es:

$$\phi_{w_2}(s) = E \left(s^{(\beta\alpha)*X} \right)$$

por la definición (1.1) del operador, tenemos:

$$\phi_{w_2}(s) = E \left(E \left(s^{\left(\sum_{i=1}^X Y_i/X \right)} \right) \right)$$

si hacemos $Z = \left(\sum_{i=1}^X Y_i/X \right) \sim Bin(\alpha\beta, X)$, tenemos:

$$\phi_{w_2}(s) = E \left(E \left((1 - \beta\alpha) s^0 + \beta\alpha s^1 \right) \right) = E (1 - \beta\alpha + \beta\alpha s) \quad \blacksquare$$

Ahora, de los resultados de los literales *a*) y *b*) se observa que $\beta * (\alpha * X)$ tiene la misma distribución que $(\beta\alpha) * X$, con lo que se verifica la propiedad (6) del operador de refinamiento.

1.3. Modelo Autorregresivo de Valores Enteros.

1.3.1. Definición del Modelo Autorregresivo de orden 1, INAR(1).

DEFINICIÓN 1.2 Sea $\{X_t; t = \pm 1, \pm 2, \pm 3, \dots\}$ una sucesión de variables aleatorias de valores enteros no negativos, entonces el proceso INAR(1) tiene la siguiente representación:

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

donde ϵ_t es v.a. i.i.d. de media μ_ϵ y varianza σ_ϵ^2 e independiente de $\alpha * X_{t-1}$.

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

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

La distribución marginal del modelo INAR(1), se expresa:

$$X_t \stackrel{d}{=} \alpha^k * X_{t-k} + \sum_{j=0}^{k-1} \alpha^j * \epsilon_{t-j} \quad (1.1)$$

Prueba:

Usando la definición 1.2 y sustituyendo recursivamente en X_{t-i} , con $i = 1, 2, 3, \dots$ tenemos:

$$\begin{aligned} X_{t-1} &= \alpha * X_{t-2} + \epsilon_{t-1} \\ X_{t-2} &= \alpha * X_{t-3} + \epsilon_{t-2} \\ &\vdots \quad \quad \quad \vdots \end{aligned}$$

usando la propiedad (6) del operador de refinamiento, tenemos:

$$\begin{aligned}
 X_t &\stackrel{d}{=} \alpha * (\alpha * X_{t-2} + \epsilon_{t-1}) + \epsilon_t \\
 &\stackrel{d}{=} \alpha^2 * X_{t-2} + \alpha * \epsilon_{t-1} + \epsilon_t \\
 &\stackrel{d}{=} \alpha^2 * (\alpha * X_{t-3} + \epsilon_{t-2}) + \alpha * \epsilon_{t-1} + \epsilon_t \\
 &\stackrel{d}{=} \alpha^3 * X_{t-3} + \alpha^2 * \epsilon_{t-2} + \alpha * \epsilon_{t-1} + \epsilon_t \\
 &\stackrel{d}{=} \alpha^3 * (\alpha * X_{t-4} + \epsilon_{t-3}) + \alpha^2 * \epsilon_{t-2} + \alpha * \epsilon_{t-1} + \epsilon_t \\
 &\stackrel{d}{=} \alpha^4 * X_{t-4} + \alpha^3 * \epsilon_{t-3} + \alpha^2 * \epsilon_{t-2} + \alpha * \epsilon_{t-1} + \epsilon_t \\
 &\stackrel{d}{=} \dots \stackrel{d}{=} \alpha^k * X_{t-k} + \sum_{j=0}^{k-1} \alpha^j * \epsilon_{t-j} \quad \blacksquare
 \end{aligned}$$

1.3.2. Propiedades Generales del Modelo Autorregresivo de orden 1, INAR(1).

Suponiendo que $E(\epsilon_t) = \mu_\epsilon < \infty$, $V(\epsilon_t) = \sigma_\epsilon^2 < \infty$, entonces el modelo INAR(1) cumple las siguientes propiedades:

1. El Valor Esperado está dado por:

$$E(X_t) = \alpha E(X_{t-1}) + \mu_\epsilon = \alpha^t E(X_0) + \mu_\epsilon \sum_{j=0}^{t-1} \alpha^j \tag{1.2}$$

Prueba:

Aplicando valor esperado al modelo de la definición 1.2, tenemos:

$$\begin{aligned}
 E(X_t) &= E(\alpha * X_{t-1} + \epsilon_t) = E(\alpha * X_{t-1}) + E(\epsilon_t) = \alpha E(X_{t-1}) + E(\epsilon_t) \\
 &= \alpha E(X_{t-1}) + \mu_\epsilon
 \end{aligned}$$

ahora, usando la ecuación (1.1), se tiene:

$$E(X_t) = \alpha^t E(X_0) + \mu_\epsilon \sum_{j=0}^{t-1} \alpha^j \quad \blacksquare$$

2. La Varianza está dada por:

$$\begin{aligned}
 V(X_t) &= \alpha^2 V(X_{t-1}) + \alpha(1-\alpha) E(X_{t-1}) + \sigma_\epsilon^2 \\
 &= \alpha^{2t} V(X_0) + (1-\alpha) \sum_{j=1}^t \alpha^{2j-1} E(X_{t-j}) + \sigma_\epsilon^2 \sum_{j=1}^t \alpha^{2(j-1)} \tag{1.3}
 \end{aligned}$$

Prueba:

$$V(X_t) = V(\alpha * X_{t-1} + \epsilon_t) = V(\alpha * X_{t-1}) + V(\epsilon_t)$$

observemos que:

$$V(\alpha * X_{t-1}) = V(E(\alpha * X_{t-1}/X_{t-1})) + E(V(\alpha * X_{t-1}/X_{t-1}))$$

como sabemos que $((\alpha * X_{t-1})/X_{t-1}) \sim Bin(\alpha, X_{t-1})$, entonces:

$$E((\alpha * X_{t-1})/X_{t-1}) = \alpha X_{t-1}$$

$$V((\alpha * X_{t-1})/X_{t-1}) = \alpha(1 - \alpha)X_{t-1}$$

$$V(X_t) = V(\alpha X_{t-1}) + E(\alpha(1 - \alpha)X_{t-1}) + V(\epsilon_t)$$

$$V(X_t) = \alpha^2 V(X_{t-1}) + \alpha(1 - \alpha)E(X_{t-1}) + \sigma_\epsilon^2$$

Sustituyendo recursivamente en X_{t-i} con $i = 1, 2, 3, \dots$ y aplicando las propiedades anteriores de la varianza, tenemos:

$$V(X_{t-i}) = \alpha^2 V(X_{t-(i+1)}) + \alpha(1 - \alpha)E(X_{t-(i+1)}) + \sigma_\epsilon^2$$

sustituyendo este término en $V(X_t)$, tenemos:

$$\begin{aligned} V(X_t) &= \alpha^2 (\alpha^2 V(X_{t-2}) + \alpha(1 - \alpha)E(X_{t-2}) + \sigma_\epsilon^2) + \alpha(1 - \alpha)E(X_{t-1}) + \sigma_\epsilon^2 \\ &= \alpha^4 V(X_{t-2}) + \alpha^3(1 - \alpha)E(X_{t-2}) + \alpha^2 \sigma_\epsilon^2 + \alpha(1 - \alpha)E(X_{t-1}) + \sigma_\epsilon^2 \\ &= \alpha^4 V(X_{t-2}) + (1 - \alpha) (\alpha E(X_{t-1}) + \alpha^3 E(X_{t-2})) + \sigma_\epsilon^2 (1 + \alpha^2) \\ &= \alpha^4 (\alpha^2 V(X_{t-3}) + \alpha(1 - \alpha)E(X_{t-3}) + \sigma_\epsilon^2) \\ &\quad + (1 - \alpha) (\alpha E(X_{t-1}) + \alpha^3 E(X_{t-2})) + \sigma_\epsilon^2 (1 + \alpha^2) \\ &= \alpha^6 V(X_{t-3}) + (1 - \alpha) (\alpha E(X_{t-1}) + \alpha^3 E(X_{t-2}) + \alpha^5 E(X_{t-3})) \\ &\quad + \sigma_\epsilon^2 (1 + \alpha^2 + \alpha^4) \\ &= \dots = \alpha^{2t} V(X_0) + (1 - \alpha) \sum_{j=1}^t \alpha^{2j-1} E(X_{t-j}) + \sigma_\epsilon^2 \sum_{j=1}^t \alpha^{2(j-1)} \quad \blacksquare \end{aligned}$$

de la expresión anterior se observa que, para que se cumpla la estacionariedad de segundo orden se requiere que el valor inicial X_0 del proceso, verifique:

$$\begin{aligned}
 a) \quad E(X_0) &= \frac{\mu_\epsilon}{(1-\alpha)} \\
 b) \quad V(X_0) &= \frac{\mu_\epsilon\alpha + \sigma_\epsilon^2}{1-\alpha^2}
 \end{aligned}$$

sustituyendo éstas expresiones en $E(X_t)$ y $V(X_t)$, se obtienen expresiones para un proceso estacionario de segundo orden.

3. La Autocovarianza de retardo k , $\gamma(k)$ está dado por:

$$Cov(X_{t-k}, X_t) = \alpha^k V(X_{t-k}) \tag{1.4}$$

Prueba:

De la ecuación (1.1), resulta fácil verificar que:

$$\begin{aligned}
 Cov(X_{t-k}, X_t) &= Cov\left(X_{t-k}, \alpha^k * X_{t-k} + \sum_{j=0}^{k-1} \alpha^j * \epsilon_{t-j}\right) \\
 &= Cov(X_{t-k}, \alpha^k * X_{t-k}) + Cov\left(X_{t-k}, \sum_{j=0}^{k-1} \alpha^j * \epsilon_{t-j}\right)
 \end{aligned}$$

sabiendo que, la covarianza para dos v.a. independientes es cero, tenemos:

$$Cov\left(X_{t-k}, \sum_{j=0}^{k-1} \alpha^j * \epsilon_{t-j}\right) = 0$$

ahora, expresamos el primer sumando de la covarianza en terminos de valor esperado:

$$\begin{aligned}
 \gamma(k) &= Cov(X_{t-k}, X_t) = Cov(X_{t-k}, \alpha^k * X_{t-k}) \\
 &= E[(X_{t-k} - E(X_{t-k}))(\alpha^k * X_{t-k} - E(\alpha^k * X_{t-k}))] \\
 &= E[X_{t-k}(\alpha^k * X_{t-k})] - E[X_{t-k}E(\alpha^k * X_{t-k})] - E[E(X_{t-k})(\alpha^k * X_{t-k})] \\
 &\quad + E[E(X_{t-k})E(\alpha^k * X_{t-k})]
 \end{aligned}$$

Ahora, aplicando la propiedad (3) y (4) del operador de refinamiento, tenemos:

$$\begin{aligned}
 \gamma(k) &= E[X_{t-k}(\alpha^k * X_{t-k})] - E[X_{t-k}E(\alpha^k * X_{t-k})] \\
 &\quad - E[E(X_{t-k})(\alpha^k * X_{t-k})] + E[E(X_{t-k})E(\alpha^k * X_{t-k})] \\
 &= \alpha^k E(X_{t-k}^2) - 2\alpha^k (E(X_{t-k}))^2 + \alpha^k (E(X_{t-k}))^2 \\
 &= \alpha^k [E(X_{t-k}^2) - (E(X_{t-k}))^2] = \alpha^k V(X_{t-k}) = \alpha^k \gamma(0) \quad \blacksquare
 \end{aligned}$$

4. La Autocorrelación de un modelo INAR(1) es:

$$\rho(k) = \alpha^k, k = 1, 2, 3, \dots \quad (1.5)$$

En efecto, este resultado se obtiene de dividir la $Cov(X_{t-k}, X_t)$ entre la $V(X_t)$. ■

La ecuación (1.5) muestra que la función de autocorrelación $\rho(k)$, decae de manera exponencial con retardo k y es idéntica en forma a la ecuación de Yule-Walker para un modelo AR(1) estándar.

Se observa que las propiedades de un proceso INAR(1) coinciden con las de un modelo AR(1) estándar, con la diferencia que el proceso INAR(1) tiene $E(X_t) = \mu$ y $V(X_t) = \sigma^2$ de una distribución discreta.

1.3.3. Definición del Modelo Autorregresivo de orden 2, INAR(2).

Alzaid y Al-Osh (1990), estudian el proceso INAR(2) y consideran que es utilizado cuando el proceso de conteo X_t , no solo depende de X_{t-1} , sino además de X_{t-2} , es decir, X_t depende de lo ocurrido dos tiempos anteriores a él.

DEFINICIÓN 1.3 Sea $\{X_t; t = \pm 1, \pm 2, \pm 3, \dots\}$ una sucesión de v.a. de valores enteros no negativos, tal que:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \epsilon_t \text{ para } t = 0, \pm 1, \pm 2, \dots$$

donde ϵ_t , $t = 0, \pm 1, \pm 2, \dots$ es una sucesión de v.a. i.i.d. de valores enteros no negativos con alguna distribución discreta, de media $\mu_\epsilon < \infty$ y varianza $\sigma_\epsilon^2 < \infty$, $(\alpha_i, i = 1, 2)$ son constantes no negativas tal que $\alpha_1 + \alpha_2 < 1$ y la distribución del vector $((\alpha_1 * X_t, \alpha_2 * X_t) / X_t = x)$ es multibinomial con parámetros (x, α_1, α_2) y es independiente de la historia pasada del proceso. Es decir, dada $X_t = x$ la v.a. $\alpha_i * X_t$ es independiente de X_{t-k} y sus sobrevivientes $\alpha_j * X_{t-k}$ para $i, j = 1, 2$, y $k > 0$.

1.3.4. Propiedades Generales del Modelo Autorregresivo de orden 2, INAR(2).

Estudiaremos las propiedades del proceso INAR(2), donde Alzaid y Al-Osh (1990) asumen que dado $X_t = x$ la v.a. $\alpha_i * X_t$ es independiente de la historia pasada del proceso $\{X_{t-k}\}$ y de $\alpha_j * X_{t-k}$ para $i, j = 1, 2$ y $k \geq 1$. Ahora, podemos ver de la definición 1.3 que:

1. El Valor Esperado está dado por:

$$\mu = \frac{\mu_\epsilon}{1 - \alpha_1 - \alpha_2} \quad (1.6)$$

Prueba:

$$\begin{aligned} E(X_t) &= E(\alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \epsilon_t) \\ &= E(\alpha_1 * X_{t-1}) + E(\alpha_2 * X_{t-2}) + E(\epsilon_t) \\ &= \alpha_1 E(X_{t-1}) + \alpha_2 E(X_{t-2}) + E(\epsilon_t) \end{aligned}$$

Suponiendo que el modelo es estacionario, tenemos:

$$\begin{aligned}
 E(X_t) &= E(X_{t-1}) = E(X_{t-2}) = \mu \\
 E(X_t) &= \alpha_1 E(X_{t-1}) + \alpha_2 E(X_{t-2}) + E(\epsilon_t) \\
 \mu &= \alpha_1 \mu + \alpha_2 \mu + \mu_\epsilon \Rightarrow \mu(1 - \alpha_1 - \alpha_2) = \mu_\epsilon \\
 \mu &= \frac{\mu_\epsilon}{(1 - \alpha_1 - \alpha_2)} \quad \blacksquare
 \end{aligned}$$

2. La Autocovarianza está dada por:

$$\gamma(k) = \alpha_1 \gamma(k-1) + \alpha_2 \gamma(k-2) + \delta_k(0) \sigma_\epsilon^2 \quad (1.7)$$

Prueba:

Usando la definición 1.3, la función de autocovarianza de retardo k podemos expresarla como:

$$\begin{aligned}
 Cov(X_{t-k}, X_t) &= Cov(X_{t-k}, \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \epsilon_t) \\
 &= Cov(X_{t-k}, \alpha_1 * X_{t-1}) + Cov(X_{t-k}, \alpha_2 * X_{t-2}) + Cov(X_{t-k}, \epsilon_t) \\
 &= Cov(X_{t-k}, \alpha_1 * X_{t-1}) + Cov(X_{t-k}, \alpha_2 * X_{t-2}) + Cov(\epsilon_{t-k}, \epsilon_t) \\
 &= Cov(X_{t-k}, \alpha_1 * X_{t-1}) + Cov(X_{t-k}, \alpha_2 * X_{t-2}) + \delta_k(0) \sigma_\epsilon^2 \\
 &= \gamma(k-1, \alpha_1) + \gamma(k-2, \alpha_2) + \delta_k(0) \sigma_\epsilon^2 \\
 &= \alpha_1 \gamma(k-1) + \alpha_2 \gamma(k-i) + \delta_k(0) \sigma_\epsilon^2
 \end{aligned}$$

donde $\gamma(k-i, \alpha_i) \equiv Cov(X_{t-k}, \alpha_i * X_{t-i}) = \alpha_i \gamma(k-i)$ (usando los resultados de la ecuación (1.4)) para $i = 1, 2$ y $\delta_k(0) = 1$ si $k = 0$ y cero en otro caso. \blacksquare

3. La Autocorrelación está dada por:

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) \quad (1.8)$$

En efecto:

Haciendo $k = 0$ en la función de autocovarianza (ecuación (1.7)), se tiene la expresión de la varianza, luego se divide la covarianza de retardo k entre la expresión de la varianza. \blacksquare

1.3.5. Definición del Modelo Autorregresivo de orden p , INAR(p).

Alzaid y Al-Osh (1990), consideran que el proceso INAR(1) descrito en la definición 1.2 es apropiado para modelar procesos de ramificación con inmigración. Sin embargo, las realizaciones de algunos procesos de conteo $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ podrían estar relacionadas no sólo con el pasado inmediato X_{t-1} , sino también con realizaciones previas del proceso $\{X_{t-i}, t = 1, 2, 3, \dots, p\}$ para alguna constante p . En consecuencia, para modelizar tales procesos y tener en consideración estas realizaciones previas, surge la necesidad de extender el proceso INAR(1) a un proceso autorregresivo de orden p .

DEFINICIÓN 1.4 Sea $\{X_t; t = \pm 1, \pm 2, \pm 3, \dots\}$ una sucesión de v.a. de valores enteros no negativos, tal que:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \dots + \alpha_p * X_{t-p} + \epsilon_t \text{ para } t = 0, \pm 1, \pm 2, \dots$$

donde $\epsilon_t, t = 0, \pm 1, \pm 2, \dots$ es una sucesión de v.a. i.i.d. de valores enteros no negativos con alguna distribución discreta, de media $\mu_\epsilon < \infty$ y varianza $\sigma_\epsilon^2 < \infty$, $(\alpha_i, i = 1, 2, 3, \dots)$ son constantes no negativas tal que $\sum_{i=1}^p \alpha_i < 1$ y la distribución del vector $((\alpha_1 * X_t, \alpha_2 * X_t, \dots, \alpha_p * X_t) / X_t = x)$ es multibinomial con parámetros $(x, \alpha_1, \alpha_2, \dots, \alpha_p)$ y es independiente de la historia pasada del proceso. Es decir, dada $X_t = x$ la variable aleatoria $\alpha_i * X_t$ es independiente de X_{t-k} y sus sobrevivientes $\alpha_j * X_{t-k}$ para $i, j = 1, 2, 3, \dots, p$ y $k > 0$.

Podemos decir que el proceso INAR(p) definido por Alzaid y Al-Osh es similar en forma, al proceso AR(p) estándar, sin embargo, la dependencia a través del tiempo del operador ‘*’ en la Definición 1.4, hace que este proceso difiera del proceso AR(p) estándar. Para explicar este punto consideremos por simplicidad el caso $p = 2$, esto es:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \epsilon_t$$

En el proceso AR(2) clásico, X_t se obtiene por una multiplicación directa de las constantes α_1 y α_2 por X_{t-1} y X_{t-2} , respectivamente, en el tiempo t e independientemente de la estructura estocástica previa. Este no es el caso para el proceso INAR(2): para este proceso las v.a. $\alpha_1 * X_{t-2}$ y $\alpha_2 * X_{t-2}$ son dependientes y se obtienen a la vez, aunque parecen ejecutadas en tiempos diferentes, ya que son elementos de X_{t-1} y X_t , respectivamente. La estructura del

proceso INAR(2) propuesto por Alzaid y Al-Osh (1990) es quizá más comprensible cuando se hace una simulación de dicho proceso. Supongamos que se han observado X_1 y X_2 , entonces el primer paso es elegir simultáneamente $\alpha_1 * X_1$ y $\alpha_2 * X_1$ de una distribución multibinomial $(X_1, \alpha_1, \alpha_2)$ y por otro lado $\alpha_1 * X_2$ y $\alpha_2 * X_2$ de una distribución multibinomial $(X_2, \alpha_1, \alpha_2)$. Por lo tanto, $X_3 = \alpha_1 * X_2 + \alpha_2 * X_1 + \epsilon_3$ y $\alpha_2 * X_2$ está disponible para ser utilizada en la derivación de X_4 . En el segundo paso, se obtienen $\alpha_1 * X_3$ y $\alpha_2 * X_3$ de una distribución multibinomial $(X_3, \alpha_1, \alpha_2)$ y, entonces, $X_4 = \alpha_1 * X_3 + \alpha_2 * X_2 + \epsilon_4$ y $\alpha_2 * X_3$ está disponible para ser utilizada en la derivación de X_5 . Así sucesivamente, se genera la muestra (X_1, X_2, \dots, X_n) del proceso INAR(2). Las suposiciones dadas en la definición 1.4, evitan contar más de una vez un elemento en el modelo.

1.3.6. Propiedades Generales del Modelo Autorregresivo de orden p , INAR(p).

Estudiaremos las propiedades del proceso INAR(p), donde Alzaid y Al-Osh (1990) asumen que dado $X_t = x$ la v.a. $\alpha_i * X_t$ es independiente de la historia pasada del proceso $\{X_{t-k}\}$ y de $\alpha_j * X_{t-k}$ para $i, j = 1, 2, \dots, p$ y $k \geq 1$. Ahora, podemos ver de la definición 1.4 que:

1. El Valor Esperado está dado por:

$$\mu = \mu_\epsilon \left[1 - \sum_{i=1}^p \alpha_i \right]^{-1} \quad (1.9)$$

Prueba:

Aplicando valor esperado a la definición 1.4, tenemos:

$$\begin{aligned} E(X_t) &= E(\alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \dots + \alpha_p * X_{t-p} + \epsilon_t) \\ &= E(\alpha_1 * X_{t-1}) + E(\alpha_2 * X_{t-2}) + \dots + E(\alpha_p * X_{t-p}) + E(\epsilon_t) \end{aligned}$$

usando la propiedad (3) del operador de refinamiento:

$$E(X_t) = \alpha_1 E(X_{t-1}) + \alpha_2 E(X_{t-2}) + \dots + \alpha_p E(X_{t-p}) + E(\epsilon_t)$$

por la condición de estacionariedad del proceso, se tiene:

$$\begin{aligned} E(X_t) &= \alpha_1 E(X_{t-1}) + \alpha_2 E(X_{t-2}) + \dots + \alpha_p E(X_{t-p}) + E(\epsilon_t) \\ \mu &= \alpha_1 \mu + \alpha_2 \mu + \dots + \alpha_p \mu + \mu_\epsilon \end{aligned}$$

$$\begin{aligned} \mu(1 - \alpha_1 - \alpha_2 - \dots - \alpha_p) &= \mu_\epsilon \\ \mu &= \frac{\mu_\epsilon}{(1 - \alpha_1 - \alpha_2 - \dots - \alpha_p)} \\ \mu &= \mu_\epsilon \left[1 - \sum_{i=1}^p \alpha_i \right]^{-1} \quad \blacksquare \end{aligned}$$

2. La Función de Autocovarianza está dada por:

$$\gamma(k) = \sum_{i=1}^p \alpha_i \gamma(k - i) + \delta_k(0) \sigma_\epsilon^2 \tag{1.10}$$

Prueba:

Sustituyendo la expresión de X_t dada en la definición 1.4, la función de autocovarianza de retardo k podemos expresarla como:

$$\begin{aligned} \gamma(k) = Cov(X_{t-k}, X_t) &= Cov(X_{t-k}, \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \dots + \alpha_p * X_{t-p} + \epsilon_t) \\ &= \sum_{i=1}^p Cov(X_{t-k}, \alpha_i * X_{t-i}) + Cov(X_{t-k}, \epsilon_t) \\ &= \sum_{i=1}^p Cov(X_{t-k}, \alpha_i * X_{t-i}) + Cov(\epsilon_{t-k}, \epsilon_t) \\ &= \sum_{i=1}^p Cov(X_{t-k}, \alpha_i * X_{t-i}) + \delta_k(0) \sigma_\epsilon^2 \\ &= \sum_{i=1}^p \gamma(k - i, \alpha_i) + \delta_k(0) \sigma_\epsilon^2 \\ &= \sum_{i=1}^p \alpha_i \gamma(k - i) + \delta_k(0) \sigma_\epsilon^2 \end{aligned}$$

donde $\gamma(k - i, \alpha_i) \equiv Cov(X_{t-k}, \alpha_i * X_{t-i}) = \alpha_i \gamma(k - i)$ (usando los resultados de la ecuación (1.4)) y $\delta_k(0) = 1$ si $k = 0$ y cero en otro caso. \blacksquare

3. La Función de Autocorrelación está dada por:

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \dots + \alpha_p \rho(k-p) \quad (1.11)$$

En efecto:

Haciendo $k = 0$ en la función de autocovarianza (ecuación (1.10)), se tiene la expresión de la varianza, por lo que tenemos:

$$\frac{\gamma(k)}{\gamma(0)} = \alpha_1 \frac{\gamma(k-1)}{\gamma(0)} + \alpha_2 \frac{\gamma(k-2)}{\gamma(0)} + \dots + \alpha_p \frac{\gamma(k-p)}{\gamma(0)}$$

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \dots + \alpha_p \rho(k-p) \quad \blacksquare$$

1.3.7. Estacionariedad.

A continuación se presentan, las condiciones de estacionariedad y la distribución límite del proceso INAR(p). Para ello, se define la sucesión de pesos $\{w_i, i = 0, 1, 2, \dots\}$ tal que:

$$\begin{aligned} w_0 &= 1 \\ w_j &= \sum_{i=1}^{\min(j,p)} \alpha_i w_{j-i} \end{aligned}$$

Ahora, si las raíces del polinomio

$$Z^p - \alpha_1 Z^{p-1} - \dots - \alpha_{p-1} Z - \alpha_p = 0 \text{ con } \alpha_p \neq 0 \quad (1.12)$$

tiene módulo menor que 1, entonces existe $\lambda, 0 < \lambda < 1$ tal que

$$0 \leq w_j \leq c\lambda^j, \quad j = 0, 1, 2, \dots$$

para alguna constante c .

Si las raíces de la ecuación (1.12) tiene módulo menor que 1 y $\sum_{j=0}^{\infty} (j+1) r_j < \infty$ con $r_j =$

$\sum_{k=j+1}^{\infty} P(\epsilon_t = k)$, entonces el proceso INAR(p) con parámetros $\alpha_1, \alpha_2, \dots, \alpha_p$ tal que $\alpha_p > 0$ es estacionario y tiene distribución límite, con f.g.p. dada por:

$$G(s) = \prod_{i=0}^{\infty} G_{\epsilon}(1 - w_i + w_i s)$$

donde $|s| \leq 1$, $G(s)$ y $G_\epsilon(z)$ denotan la f.g.p. de X_t y ϵ_t , respectivamente.

Prueba:

Las sucesiones de pesos para un modelo INAR(p), están dadas por:

$$\begin{aligned}
 w_0 &= 1 \\
 w_1 &= \alpha_1 \\
 w_2 &= \alpha_1 w_1 + \alpha_2 w_0 \\
 w_3 &= \alpha_1 w_2 + \alpha_2 w_1 + \alpha_3 w_0 \\
 w_4 &= \alpha_1 w_3 + \alpha_2 w_2 + \alpha_3 w_1 + \alpha_4 w_0 \\
 w_4 &= \alpha_1 w_4 + \alpha_2 w_3 + \alpha_3 w_2 + \alpha_4 w_1 + \alpha_5 w_0 \\
 &\vdots
 \end{aligned}$$

Recordemos que el modelo INAR(p), está dado por:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \dots + \alpha_p * X_{t-p} + \epsilon_t$$

ahora, de este modelo, y sabiendo que solo se conocen las observaciones de 1 hasta t , podemos generar los X_i de manera recursiva y usando el resultado de la ecuación (1.1), se tiene:

$$\begin{aligned}
 X_1 &= \epsilon_1 \\
 X_2 &= \alpha_1 * X_1 + \epsilon_2 = \alpha_1 * \epsilon_1 + \epsilon_2 \\
 X_3 &= \alpha_1 * X_2 + \alpha_2 * X_1 + \epsilon_3 \\
 &\stackrel{d}{=} \alpha_1 * (\alpha_1 * \epsilon_1 + \epsilon_2) + \alpha_2 * \epsilon_1 + \epsilon_3 \stackrel{d}{=} (\alpha_1^2 + \alpha_2) * \epsilon_1 + \alpha_1 * \epsilon_2 + \epsilon_3 \\
 X_4 &= \alpha_1 * X_3 + \alpha_2 * X_2 + \alpha_3 * X_1 + \epsilon_4 \\
 &\stackrel{d}{=} \alpha_1 * ((\alpha_1^2 + \alpha_2) * \epsilon_1 + \alpha_1 * \epsilon_2 + \epsilon_3) + \alpha_2 * (\alpha_1 * \epsilon_1 + \epsilon_2) + \alpha_3 * \epsilon_1 + \epsilon_4 \\
 &\stackrel{d}{=} (\alpha_1^3 + 2\alpha_1\alpha_2 + \alpha_3) * \epsilon_1 + (\alpha_1^2 + \alpha_2) * \epsilon_2 + \alpha_1 * \epsilon_3 + \epsilon_4 \\
 &\vdots
 \end{aligned}$$

se observa que $\{w_j\}_{j=0}^{\infty}$ son los coeficientes de la representación en medias móviles para el proceso INAR(p) es decir:

$$\begin{aligned} X_1 &\stackrel{d}{=} w_0 * \epsilon_1 \\ X_2 &\stackrel{d}{=} w_1 * \epsilon_1 + w_0 * \epsilon_2 \\ X_3 &\stackrel{d}{=} w_2 * \epsilon_1 + w_1 * \epsilon_2 + w_0 * \epsilon_3 \\ X_4 &\stackrel{d}{=} w_3 * \epsilon_1 + w_2 * \epsilon_2 + w_1 * \epsilon_3 + w_0 * \epsilon_4 \\ &\vdots \\ X_i &\stackrel{d}{=} w_{i-1} * \epsilon_1 + w_{i-2} * \epsilon_2 + \dots + w_0 * \epsilon_i \\ &\vdots \end{aligned}$$

Entonces, la f.g.p. está dada por:

$$G(s) = E(s^{X_t}) = E(s^{w_0 * \epsilon_t + w_1 * \epsilon_{t-1} + w_2 * \epsilon_{t-2} + \dots}) = E(s^{w_0 * \epsilon_t} s^{w_1 * \epsilon_{t-1}} s^{w_2 * \epsilon_{t-2}} s^{\dots})$$

dado que son sucesos independientes, se puede escribir:

$$\begin{aligned} G(s) &= \prod_{j=0}^{\infty} E(s^{w_j * \epsilon_{t-j}}) = \prod_{j=0}^{\infty} E(E(s^{w_j * \epsilon_{t-j}} / \epsilon_{t-j})) = \prod_{j=0}^{\infty} G_{\epsilon}((1 - w_j)s^0 + w_j s^1) \\ &= \prod_{j=0}^{\infty} G_{\epsilon}(1 - w_j + w_j s) \quad \blacksquare \end{aligned}$$

1.4. Modelo de Media Móvil de Valores Enteros.

1.4.1. Definición del Modelo de Media Móvil de orden 1, INMA(1).

Se discutirá una secuencia estacionaria de v.a. $\{X_t\}$ que está formada por una secuencia $\{\epsilon_t\}$ de v.a. enteras no negativas i.i.d. enfocándonos en los casos en el que dos valores próximos están correlados, el tipo de modelo de primer orden de media móvil está relacionado con el proceso INAR(1) (ver definición 1.2).

DEFINICIÓN 1.5 Sea $\{X_t; t = \pm 1, \pm 2, \pm 3, \dots\}$ una sucesión de v.a. de valores enteros no negativos, entonces el proceso INMA(1) admite la forma:

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

donde ϵ_t es v.a. i.i.d. de media μ_ϵ y varianza σ_ϵ^2 .

El modelo es similar al proceso estándar MA(1) en el que la multiplicación escalar es reemplazada por la operación “ $\beta*$ ”, de la misma manera que se definió en el proceso INAR(1). Tal reemplazo permite más variabilidad en el proceso. El incremento en la varianza del proceso es porque $\beta * \epsilon_t$ condicionado por ϵ_t es todavía una v.a. y en consecuencia tiene varianza asociada. De hecho, para un valor de ϵ_t dado, $\beta * \epsilon_t$ tiene una distribución binomial con parámetros (ϵ_t, β) . Puesto que el número de conteos en un proceso, durante un cierto período de tiempo puede ser considerado como la suma de v.a. binarias.

1.4.2. Propiedades Generales del Modelo de Media Móvil de orden 1, INMA(1).

Estudiando las propiedades del proceso INMA(1), debemos hacer uso de la independencia de las v.a. $\beta * \epsilon_{t-i}$ y $\beta * \epsilon_{t-j}$ que son v.a. independientes para $i \neq j$; el cual sigue nuestras asunciones concernientes a la definición 1.5:

1. El Valor Esperado está dado por:

$$E(X_t) = (1 + \beta) \mu_\epsilon \tag{1.13}$$

Prueba:

$$E(X_t) = E(\beta * \epsilon_{t-1} + \epsilon_t) = E(\beta * \epsilon_{t-1}) + E(\epsilon_t)$$

aplicando la propiedad (3) del operador de refinamiento y suponiendo que el valor esperado de $E(\epsilon_t) = \mu_\epsilon$, tenemos:

$$E(X_t) = \beta E(\epsilon_{t-1}) + E(\epsilon_t) = \beta \mu_\epsilon + \mu_\epsilon = (1 + \beta) \mu_\epsilon \quad \blacksquare$$

2. La Varianza está dado por:

$$V(X_t) = \beta(1 - \beta) \mu_\epsilon + (1 + \beta^2) \sigma_\epsilon^2 \quad (1.14)$$

Prueba:

$$V(X_t) = V(\beta * \epsilon_{t-1} + \epsilon_t) = V(\beta * \epsilon_{t-1}) + V(\epsilon_t)$$

De igual manera que para la varianza de un modelo INAR(1) y suponiendo que la varianza de $V(\epsilon_t) = \sigma_\epsilon^2$:

$$V(\beta * \epsilon_{t-1}) = \beta^2 V(\epsilon_{t-1}) + \beta(1 - \beta) E(\epsilon_{t-1})$$

por lo que la varianza se expresa:

$$\begin{aligned} V(\epsilon_t) &= \beta^2 V(\epsilon_{t-1}) + \beta(1 - \beta) E(\epsilon_{t-1}) + V(\epsilon_t) = \beta^2 \sigma_\epsilon^2 + \beta(1 - \beta) \mu_\epsilon + \sigma_\epsilon^2 \\ &= \beta(1 - \beta) \mu_\epsilon + (1 + \beta^2) \sigma_\epsilon^2 \quad \blacksquare \end{aligned}$$

3. La Covarianza está dado por:

$$Cov(X_{t-1}, X_t) = Cov(\epsilon_{t-1}, \beta * \epsilon_{t-1}) = \beta \sigma_\epsilon^2 \quad (1.15)$$

Prueba:

$$\begin{aligned} X_t &= \beta * \epsilon_{t-1} + \epsilon_t \\ X_{t-1} &= \beta * \epsilon_{t-2} + \epsilon_{t-1} \end{aligned}$$

$$\begin{aligned}
 Cov(X_t, X_{t-1}) &= Cov(\beta * \epsilon_{t-1} + \epsilon_t, \beta * \epsilon_{t-2} + \epsilon_{t-1}) \\
 &= Cov(\beta * \epsilon_{t-1}, \beta * \epsilon_{t-2} + \epsilon_{t-1}) + Cov(\epsilon_t, \beta * \epsilon_{t-2} + \epsilon_{t-1}) \\
 &= Cov(\beta * \epsilon_{t-1}, \epsilon_{t-1})
 \end{aligned}$$

dado que la covarianza para retardos diferentes es cero.

Esta última expresión es similar a la encontrada para la covarianza de retardo k de un modelo INAR(1) y suponiendo que $V(\epsilon_{t-1}) = \sigma_\epsilon^2$:

$$Cov(X_{t-1}, X_t) = Cov(\epsilon_{t-1}, \beta * \epsilon_{t-1}) = \beta V(\epsilon_{t-1}) = \beta \sigma_\epsilon^2 \quad \blacksquare$$

4. La Autocorrelación está dado por:

$$\rho(k) = \begin{cases} \frac{\beta \sigma_\epsilon^2}{\beta(1-\beta)\mu_\epsilon + (1+\beta^2)\sigma_\epsilon^2} & \text{si } k = 1 \\ 0 & \text{si } k > 1 \end{cases} \quad (1.16)$$

En efecto:

Para $k = 1$, el resultado se obtiene de dividir la ecuación (1.15) entre la ecuación (1.14), mientras que para $k > 1$, el resultado se observa de la propiedad de la covarianza. \blacksquare

En el proceso INMA(1), las correlaciones de alto orden de la serie son cero. El resultado de la ecuación (1.16) implica que la primera correlación de la serie es no negativa y limitada por 1/2 (Valor que $\rho(1)$ toma cuando β tiende a uno), además, de un modelo MA(1) estándar, sabemos que $\rho(1) \in [-1, 0]$ si $\beta \in [-1, 0]$ y $\rho(1) \in [0, 1]$ si $\beta \in [0, 1]$, entonces, la restricción en $\rho(1) = 1$ para el proceso INMA(1) es una consecuencia de no negatividad del modelo.

Un simple aspecto del proceso INMA(1) es que para $\beta = 0$, la secuencia de X_t reducida a una secuencia independiente de v.a. que tiene la distribución de ϵ_t . También, el máximo valor de ρ ocurre cuando $\beta = 1$. Esta propiedad del proceso es semejante a la de un proceso MA(1) estándar.

1.4.3. Función Generatriz de Probabilidad del Modelo INMA(1).

La función generatriz de probabilidad (f.g.p.) está dada por:

$$\phi_{X_t}(s) = E(s^{X_t}) = E((1 - \beta + \beta s)^{\epsilon_{t-1}} s^{\epsilon_t}) = \psi(1 - \beta + \beta s) \psi(s), \quad |s| \leq 1 \quad (1.17)$$

donde $\psi(s)$ es la f.g.p. de ϵ_t .

Prueba:

$$\begin{aligned} \phi_{X_t}(s) &= E(s^{X_t}) = E(s^{\beta * \epsilon_{t-1} + \epsilon_t}) = E(s^{\beta * \epsilon_{t-1}} s^{\epsilon_t}) = E\left(s^{\left(\sum_{i=1}^{\epsilon_{t-1}} Y_i\right)} s^{\epsilon_t}\right) \\ &= E\left(E\left(s^{\left(\sum_{i=1}^{\epsilon_{t-1}} Y_i / \epsilon_{t-1}\right)}\right) s^{\epsilon_t}\right) = E\left(E\left(s^{\left(\sum_{i=1}^{\epsilon_{t-1}} Y_i / \epsilon_{t-1}\right)}\right)\right) E(s^{\epsilon_t}) \end{aligned}$$

dado que ϵ_{t-1} y ϵ_t son mutuamente independientes y sabemos que:

$$((\beta * \epsilon_{t-1}) / \epsilon_{t-1}) \sim \text{Bin}(\epsilon_{t-1}, \beta)$$

$$\begin{aligned} \phi_{X_t}(s) &= E\left(\left((1 - \beta) s^0 + \beta s^1\right)^{\epsilon_{t-1}}\right) E(s^{\epsilon_t}) = E\left((1 - \beta + \beta s)^{\epsilon_{t-1}}\right) E(s^{\epsilon_t}) \\ &= \psi(1 - \beta + \beta s) \psi(s) \end{aligned}$$

En general, la f.g.p. para el vector (X_1, X_2, \dots, X_t) es:

$$\phi_t(s_1, s_2, \dots, s_t) = \psi(s_t) \prod_{i=1}^t \psi(s_{i-1} (1 - \beta + \beta s_i)) \quad \text{para } |s_i| \leq 1, \quad s_0 = 1 \quad (1.18)$$

Prueba:

$$\phi_t(s_1, s_2, \dots, s_t) = E(s_1^{X_1} s_2^{X_2} \dots s_t^{X_t})$$

sustituyendo $X_t = \beta * \epsilon_{t-1} + \epsilon_t$, tenemos:

$$\begin{aligned} \phi_t(s_1, s_2, \dots, s_t) &= E(s_1^{X_1} s_2^{X_2} \dots s_t^{X_t}) = E\left(s_1^{X_1} s_2^{X_2} \dots s_t^{\beta * \epsilon_{t-1} + \epsilon_t}\right) \\ &= E\left(s_1^{X_1} s_2^{X_2} \dots s_t^{\beta * \epsilon_{t-1}} s_t^{\epsilon_t}\right) \end{aligned}$$

dado que ϵ_t no aparece en $\{X_1, X_2, \dots, X_{t-1}\}$, se vuelven independientes:

$$\begin{aligned}
\phi_t(s_1, s_2, \dots, s_t) &= E\left(s_1^{X_1} s_2^{X_2} \dots s_t^{\beta * \epsilon_{t-1}} s_t^{\epsilon_t}\right) = E\left(s_1^{X_1} s_2^{X_2} \dots s_t^{\beta * \epsilon_{t-1}}\right) E\left(s_t^{\epsilon_t}\right) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{X_{t-1}} s_t^{\beta * \epsilon_{t-1}}\right) E\left(s_t^{\epsilon_t}\right) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2} + \epsilon_{t-1}} s_t^{\beta * \epsilon_{t-1}}\right) E\left(s_t^{\epsilon_t}\right) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) E\left(s_{t-1}^{\epsilon_{t-1}} s_t^{\beta * \epsilon_{t-1}}\right) E\left(s_t^{\epsilon_t}\right) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) E\left(E\left(s_{t-1}^{\epsilon_{t-1}} s_t^{\beta * \epsilon_{t-1}} / \epsilon_{t-1}\right)\right) E\left(s_t^{\epsilon_t}\right)
\end{aligned}$$

dado que $s_{t-1}^{\epsilon_{t-1}}$ y $s_t^{\beta * \epsilon_{t-1}}$ son independientes dado ϵ_{t-1} :

$$\phi_t(s_1, s_2, \dots, s_t) = E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) E\left(E\left(s_{t-1}^{\epsilon_{t-1}} / \epsilon_{t-1}\right) E\left(s_t^{\beta * \epsilon_{t-1}} / \epsilon_{t-1}\right)\right) E\left(s_t^{\epsilon_t}\right)$$

sabemos que $((\beta * \epsilon_{t-1}) / \epsilon_{t-1}) \sim Bin(\epsilon_{t-1}, \beta)$:

$$\begin{aligned}
\phi_t(s_1, s_2, \dots, s_t) &= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) \phi_{\epsilon_{t-1}}(s_{t-1}) \\
&\times E\left(\left((1 - \beta) s_t^0 + \beta s_t^1\right)^{\epsilon_{t-1}}\right) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) \phi_{\epsilon_{t-1}}(s_{t-1}) \\
&\times E\left(\left(1 - \beta + \beta s_t\right)^{\epsilon_{t-1}}\right) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-1}^{\beta * \epsilon_{t-2}}\right) \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t)
\end{aligned}$$

ahora, sustituyendo $X_{t-2} = \beta * \epsilon_{t-3} + \epsilon_{t-2}$, se tiene:

$$\begin{aligned}
\phi_t(s_1, s_2, \dots, s_t) &= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta * \epsilon_{t-3} + \epsilon_{t-2}} s_{t-1}^{\beta * \epsilon_{t-2}}\right) \phi_{\epsilon_{t-1}}(s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta * \epsilon_{t-3}} s_{t-2}^{\epsilon_{t-2}} s_{t-1}^{\beta * \epsilon_{t-2}}\right) \phi_{\epsilon_{t-1}}(s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta * \epsilon_{t-3}}\right) E\left(s_{t-2}^{\epsilon_{t-2}} s_{t-1}^{\beta * \epsilon_{t-2}}\right) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t)
\end{aligned}$$

dado que ϵ_{t-2} no aparece en $\{X_1, X_2, \dots, X_{t-3}\}$, se vuelven independientes y tenemos:

$$\begin{aligned}
\phi_t(s_1, s_2, \dots, s_t) &= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta^* \epsilon_{t-3}}\right) E\left(E\left(s_{t-2}^{\epsilon_{t-2}} s_{t-1}^{\beta^* \epsilon_{t-2}} / \epsilon_{t-2}\right)\right) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta^* \epsilon_{t-3}}\right) E\left(s_{t-2}^{\epsilon_{t-2}} / \epsilon_{t-2}\right) E\left(s_{t-1}^{\beta^* \epsilon_{t-2}} / \epsilon_{t-2}\right) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-2}^{\beta^* \epsilon_{t-3}}\right) \phi_{\epsilon_{t-2}}(s_{t-2}) \phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= \\
&\vdots \\
&= E\left(s_1^{X_1} s_2^{X_2} \dots s_{t-3}^{\beta^* \epsilon_{t-4}}\right) \\
&\times \phi_{\epsilon_{t-3}}(s_{t-3}) \phi_{\epsilon_{t-3}}(1 - \beta + \beta s_{t-2}) \phi_{\epsilon_{t-2}}(s_{t-2}) \phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t)
\end{aligned}$$

al sustituir recursivamente para X_{t-i} para $i = 4, 5, \dots$, se tiene:

$$\begin{aligned}
\phi_t(s_1, s_2, \dots, s_t) &= E\left(s_1^{X_1} s_2^{\beta^* \epsilon_1}\right) \phi_{\epsilon_2}(s_2) \phi_{\epsilon_2}(1 - \beta + \beta s_3) \phi_{\epsilon_3}(s_3) \phi_{\epsilon_3}(1 - \beta + \beta s_4) \\
&\times \dots \times \phi_{\epsilon_{t-3}}(s_{t-3}) \phi_{\epsilon_{t-3}}(1 - \beta + \beta s_{t-2}) \\
&\times \phi_{\epsilon_{t-2}}(s_{t-2}) \phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= E\left(s_1^{\beta^* \epsilon_0 + \epsilon_1} s_2^{\beta^* \epsilon_1}\right) \phi_{\epsilon_2}(s_2) \phi_{\epsilon_2}(1 - \beta + \beta s_3) \phi_{\epsilon_3}(s_3) \phi_{\epsilon_3}(1 - \beta + \beta s_4) \\
&\times \dots \times \phi_{\epsilon_{t-3}}(s_{t-3}) \phi_{\epsilon_{t-3}}(1 - \beta + \beta s_{t-2}) \phi_{\epsilon_{t-2}}(s_{t-2}) \phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t) \\
&= \phi_{\epsilon_0}(1 - \beta + \beta s_1) \phi_{\epsilon_1}(s_1) \phi_{\epsilon_1}(1 - \beta + \beta s_2) \\
&\times \phi_{\epsilon_2}(s_2) \phi_{\epsilon_2}(1 - \beta + \beta s_3) \phi_{\epsilon_3}(s_3) \phi_{\epsilon_3}(1 - \beta + \beta s_4) \\
&\times \dots \times \phi_{\epsilon_{t-3}}(s_{t-3}) \phi_{\epsilon_{t-3}}(1 - \beta + \beta s_{t-2}) \phi_{\epsilon_{t-2}}(s_{t-2}) \phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \\
&\times \phi_{\epsilon_{t-1}}(s_{t-1}) \phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t) \phi_{\epsilon_t}(s_t)
\end{aligned}$$

por definición sabemos que $s_0 = 1$, entonces $\phi_{\epsilon_0}(s_0) = E((s_0)^{\epsilon_0}) = 1$, la expresión no se altera si hacemos:

$$\begin{aligned}
 \phi_t(s_1, s_2, \dots, s_t) &= \phi_{\epsilon_0}(s_0)\phi_{\epsilon_0}(1 - \beta + \beta s_1)\phi_{\epsilon_1}(s_1)\phi_{\epsilon_1}(1 - \beta + \beta s_2) \\
 &\times \phi_{\epsilon_2}(s_2)\phi_{\epsilon_2}(1 - \beta + \beta s_3)\phi_{\epsilon_3}(s_3)\phi_{\epsilon_3}(1 - \beta + \beta s_4) \\
 &\times \\
 &\vdots \\
 &\times \phi_{\epsilon_{t-3}}(s_{t-3})\phi_{\epsilon_{t-3}}(1 - \beta + \beta s_{t-2})\phi_{\epsilon_{t-2}}(s_{t-2})\phi_{\epsilon_{t-2}}(1 - \beta + \beta s_{t-1}) \\
 &\times \phi_{\epsilon_{t-1}}(s_{t-1})\phi_{\epsilon_{t-1}}(1 - \beta + \beta s_t)\phi_{\epsilon_t}(s_t) \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t \phi_{\epsilon_{i-1}}(s_{i-1})\phi_{\epsilon_{i-1}}(1 - \beta + \beta s_i) \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t \phi_{\epsilon_{i-1}}(1 - \beta + \beta s_i)\phi_{\epsilon_{i-1}}(s_{i-1}) \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t E((1 - \beta + \beta s_i)^{\epsilon_{i-1}})E((s_{i-1})^{\epsilon_{i-1}}) \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t E((1 - \beta + \beta s_i)^{\epsilon_{i-1}}(s_{i-1})^{\epsilon_{i-1}}) \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t E[((1 - \beta + \beta s_i)s_{i-1})^{\epsilon_{i-1}}] \\
 &= \phi_{\epsilon_t}(s_t)\prod_{i=1}^t \phi_{\epsilon_{i-1}}(s_{i-1}(1 - \beta + \beta s_i)) \\
 &= \psi(s_t)\prod_{i=1}^t \psi(s_{i-1}(1 - \beta + \beta s_i)) \quad \blacksquare
 \end{aligned}$$

1.4.4. Aspectos Básicos de un Proceso INMA(1) de Poisson.

Suponemos ahora que ϵ_t tiene una distribución de Poisson con media λ , entonces la f.g.p. está dada por:

$$\psi(s) = e^{-\lambda(1-s)}$$

Ahora, de la ecuación (1.17) se tiene:

$$\psi_{X_t}(s) = e^{-\lambda(1+\beta)(1-s)} \quad (1.19)$$

Prueba:

$$\begin{aligned} \psi_{X_t}(s) &= \psi(1 - \beta + \beta s) \psi(s) = e^{-\lambda(1-(1-\beta+\beta s))} e^{-\lambda(1-s)} = e^{-\lambda(\beta-\beta s)} e^{-\lambda(1-s)} \\ &= e^{-\lambda\beta(1-s)} e^{-\lambda(1-s)} = e^{-\lambda\beta(1-s)-\lambda(1-s)} = e^{-\lambda(1+\beta)(1-s)} \quad \blacksquare \end{aligned}$$

Esto muestra que X_t tiene distribución de Poisson con media $\lambda(1 + \beta)$, solo la marginal de Poisson tiene esta propiedad de descomposición.

Resulta de interés estudiar la distribución del total de ocurrencias durante el tiempo $[0, k]$ el cual es denotado por $T_k = \sum_{i=1}^k X_i$. Considerando los puntos de la f.g.p. de $\{X_1, X_2, \dots, X_k\}$, usando la ecuación (1.18), se tiene:

$$\psi_{X_k}(s_1, s_2, \dots, s_k) = e^{-\lambda \left[k + \beta - (1 - \beta) \sum_{i=1}^k s_i - \beta (s_1 + s_k) - \beta \sum_{i=1}^{k-1} s_i s_{i+1} \right]} \quad (1.20)$$

Prueba:

de la ecuación (1.18) tenemos que:

$$\begin{aligned} \phi_{X_k}(s_1, s_2, \dots, s_k) &= \psi(s_k) \prod_{i=1}^k \psi(s_{i-1} (1 - \beta + \beta s_i)) \\ &= e^{-\lambda(1-s_k)} \prod_{i=1}^k e^{-\lambda(1 - s_{i-1} (1 - \beta + \beta s_i))} \end{aligned}$$

el producto de k funciones exponenciales se puede expresar como:

$$\begin{aligned}
\phi_{X_k}(s_1, \dots, s_k) &= e^{-\lambda(1-s_k)} e^{-\lambda \sum_{i=1}^k (1-s_{i-1}(1-\beta+\beta s_i))} \\
&= e^{-\lambda \left[(1-s_k) + \sum_{i=1}^k (1-s_{i-1}(1-\beta+\beta s_i)) \right]} \\
&= e^{-\lambda \left[(1-s_k) + \sum_{i=1}^k (1-s_{i-1} + \beta s_{i-1} - \beta s_i s_{i-1}) \right]} \\
&= e^{-\lambda \left[1-s_k + k - (1-\beta) \sum_{i=1}^k s_{i-1} - \beta \sum_{i=1}^k s_i s_{i-1} \right]} \\
&= e^{-\lambda \left[1-s_k + k - (1-\beta) s_0 - (1-\beta) \sum_{i=1}^{k-1} s_i - \beta \sum_{i=1}^{k-1} s_i s_{i+1} - \beta s_1 s_0 \right]}
\end{aligned}$$

sabemos que $s_0 = 1$ y además, sumamos y restamos βs_k en la expresión anterior:

$$\begin{aligned}
\phi_{X_k}(s_1, \dots, s_k) &= e^{-\lambda \left[1-s_k + k - (1-\beta) s_0 - (1-\beta) \sum_{i=1}^{k-1} s_i - \beta \sum_{i=1}^{k-1} s_i s_{i+1} - \beta s_1 s_0 \right]} \\
&= e^{-\lambda \left[1-s_k + k - 1 + \beta - (1-\beta) \sum_{i=1}^{k-1} s_i - \beta s_k + \beta s_k - \beta \sum_{i=1}^{k-1} s_i s_{i+1} - \beta s_1 \right]} \\
&= e^{-\lambda \left[k + \beta - (1-\beta) s_k - (1-\beta) \sum_{i=1}^{k-1} s_i - \beta s_k - \beta s_1 - \beta \sum_{i=1}^{k-1} s_i s_{i+1} \right]} \\
&= e^{-\lambda \left[k + \beta - (1-\beta) \sum_{i=1}^k s_i - \beta (s_1 + s_k) - \beta \sum_{i=1}^{k-1} s_i s_{i+1} \right]} \blacksquare
\end{aligned}$$

Sea ϕ_k^* la f.g.p. de T_k . Entonces usando la ecuación (1.18) tenemos:

$$\phi_k^*(s) = e^{-\lambda [(1-\beta)k + 2\beta] - \lambda\beta(k-1)(1-s^2)} \quad (1.21)$$

La última ecuación muestra que T_k se distribuye como la suma de dos v.a. independientes, $Z + 2Y$, donde Z y Y tiene una distribución de Poisson con medias $\lambda[(1-\beta)k + 2\beta]$ y $\lambda\beta(k-1)$ respectivamente.

Consecuentemente T_k no tiene distribución de Poisson. Este es un contraste con el proceso MA(1) estándar donde $\sum_{i=1}^k X_i$ tiene distribución Gaussiana.

1.4.5. Invertibilidad.

La ecuación (1.20), muestra que el proceso INMA(1) de Poisson es de tiempo invertible en el sentido que $\{X_1, X_2, \dots, X_k\}$ tienen la misma distribución que $\{X_k, X_{k-1}, \dots, X_1\}$ para todo $k \geq 2$. La invertibilidad del proceso INMA(1) es una característica de las propiedades del proceso de Poisson tal como se enuncia en el siguiente teorema:

TEOREMA 1.1 *Un proceso INMA(1) no degenerado $\{X_k\}$ es invertible si y solo si el proceso $\epsilon_t \stackrel{d}{=} \epsilon + \ell$, donde ϵ tiene una distribución de Poisson y ℓ un entero no negativo.*

La prueba de este teorema se obtendrá en dos partes:

1. El proceso INMA(1) no degenerado $\{X_t\}$ es invertible. Entonces el proceso $\epsilon_t \stackrel{d}{=} \epsilon + \ell$. Se supone que el proceso INMA(1) es invertible. Entonces usando (1.18) para $k = 2$ tenemos:

$$\begin{aligned} \psi(1 - \beta + \beta s_1) \psi(s_1(1 - \beta + \beta s_2)) \psi(s_2) &= \psi(1 - \beta + \beta s_2) \\ &\times \psi(s_2(1 - \beta + \beta s_1)) \psi(s_1) \end{aligned} \quad (1.22)$$

Sin pérdida de generalidad asumimos que $\psi(0) \neq 0$ puesto que si $\psi(0) = 0$, se puede siempre encontrar un entero ℓ tal que: $\phi^*(s) = \phi(s)/s^\ell$ con $\phi^*(0) \neq 0$, y ϕ^* satisface (1.22). Ahora si ajustamos $s_1 = 0$ y reemplazamos s_2 por s en (1.22), obtenemos:

$$\begin{aligned} \psi(1 - \beta) \psi(0) \psi(s) &= \psi(1 - \beta + \beta s) \psi(s(1 - \beta)) \psi(0) \\ \psi(s) &= \frac{\psi(1 - \beta + \beta s) \psi(s(1 - \beta))}{\psi(1 - \beta)} \end{aligned} \quad (1.23)$$

para todo $|s| \leq 1$ y $\psi(1 - \beta)$ es constante.

Derivando la ecuación (1.22) con respecto a s_1 :

$$\begin{aligned} \psi'(1 - \beta + \beta s_1) \beta \psi(s_1(1 - \beta + \beta s_2)) \psi(s_2) + \\ \psi(1 - \beta + \beta s_1) \psi'(s_1(1 - \beta + \beta s_2)) (1 - \beta + \beta s_2) \psi(s_2) = \\ \psi(1 - \beta + \beta s_2) \left[\psi'(s_2(1 - \beta + \beta s_1)) \beta s_2 \psi(s_1) + \psi(s_2(1 - \beta + \beta s_1)) \psi'(s_1) \right] \end{aligned}$$

haciendo $s_1 = 0$ y $s_2 = s$, en la expresión anterior, tenemos:

$$\begin{aligned} \left[\psi'(1 - \beta) \beta \psi(0) + \psi(1 - \beta) \psi'(0) (1 - \beta + \beta s) \right] \psi(s) \\ = \psi(1 - \beta + \beta s) \left[\psi'(s(1 - \beta)) \beta s \psi(0) + \psi(s(1 - \beta)) \psi'(0) \right] \end{aligned}$$

sustituyendo $\psi(s)$ (ecuación (1.23)), en la ecuación anterior:

$$\begin{aligned} & \left[\psi'(1-\beta)\beta\psi(0) + \psi(1-\beta)\psi'(0)(1-\beta+\beta s) \right] \frac{\psi(1-\beta+\beta s)\psi(s(1-\beta))}{\psi(1-\beta)} \\ & = \psi(1-\beta+\beta s) \left[\psi'(s(1-\beta))\beta s\psi(0) + \psi(s(1-\beta))\psi'(0) \right] \end{aligned}$$

se tiene:

$$\begin{aligned} \frac{\psi'(1-\beta)}{\psi(1-\beta)}\beta\psi(0) + \psi'(0)(1-\beta+\beta s) &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))}\beta s\psi(0) + \psi'(0) \\ \frac{\psi'(1-\beta)}{\psi(1-\beta)}\beta\psi(0) - \psi'(0)\beta(1-s) &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))}\beta s\psi(0) \\ \frac{\psi'(1-\beta)}{\psi(1-\beta)} - \frac{\psi'(0)}{\psi(0)}(1-s) &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))}s \\ \frac{\psi'(1-\beta)}{\psi(1-\beta)} - \frac{\psi'(0)}{\psi(0)} &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))}s - \frac{\psi'(0)}{\psi(0)}s \\ \left[\frac{\psi'(1-\beta)}{\psi(1-\beta)} - \frac{\psi'(0)}{\psi(0)} \right] \frac{1}{s} &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))} - \frac{\psi'(0)}{\psi(0)} \\ \frac{C_1}{s} &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))} - C_2 \\ C_2 &= \frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))} - \frac{C_1}{s} \end{aligned}$$

para C_1 y C_2 , constantes. Ahora, integrando esta última expresión con respecto a s :

$$\begin{aligned} \int C_2 ds &= \int \left(\frac{\psi'(s(1-\beta))}{\psi(s(1-\beta))} - \frac{C_1}{s} \right) ds \\ C_2 s &= \ln[\psi(s(1-\beta))] + (-C_1)\ln(s) + C_2 \\ -C_2(1-s) &= \ln[\psi(s(1-\beta))] + \ln(s)^{-C_1} \\ e^{-C_2(1-s)} &= \psi(s(1-\beta))s^{-C_1} \\ e^{-C_2(1-s)}s^{C_1} &= \psi(s(1-\beta)) \end{aligned}$$

si $C_2 = \lambda$ y $C_1 = \ell$ y resolviendo para $\psi(s)$ se tiene:

$$\psi(s) = e^{-\lambda(1-s)}$$

para alguna constante λ . Entonces, la solución general es de la forma:

$$\psi^*(s) = s^\ell e^{-\lambda(1-s)}$$

para algún entero $\ell \geq 0$.

2. El proceso $\epsilon_t \stackrel{d}{=} \epsilon + \ell$. Entonces el proceso INMA(1) no degenerado $\{X_t\}$ es invertible.

Se supone que $\epsilon_t \stackrel{d}{=} \epsilon + \ell$, entonces:

La f.g.p. de $\epsilon_t \sim P(\lambda)$ es:

$$\psi(s) = e^{-\lambda(1-s)}$$

la f.g.p. de $\epsilon + \ell$ donde $\epsilon \sim P(\lambda)$ y ℓ un entero positivo es:

$$\begin{aligned}\psi^*(s) &= E(s^{\epsilon+\ell}) = E(s^\epsilon s^\ell) = s^\ell E(s^\epsilon) \\ &= s^\ell e^{-\lambda(1-s)}\end{aligned}$$

De los literales 1) y 2) se obtiene el mismo resultado, por lo tanto, se verifica la prueba de invertibilidad de un proceso INMA(1). ■

1.4.6. Definición del Modelo de Media Móvil de orden 2, INMA(2).

Considerando la extensión del modelo INMA(1) y que fue estudiada por Al-Osh y Alzaid (1988b), se presenta el modelo INMA(2), consideran un proceso en el que la máxima vida para un elemento en el sistema es (3) unidades de tiempo. En primer lugar, se consideran α_j ($j = 1, 2$) 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 $\alpha_1 + \alpha_2 \leq 1$.

Se definen los coeficientes β_i ($i = 0, 1, 2$) como:

$$\beta_0 = 1$$

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

Los coeficientes β_i pueden interpretarse como la probabilidad que tiene un elemento de ϵ_t de ser un elemento de X_{t+i} (este elemento podría estar presente en el sistema durante instantes precedentes al $t + i$).

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

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

entonces $\beta_j = P(Y_{i,j}^t = 1)$ para $j = 1, 2$ y para todo $i = 1, 2, \dots$ y $t = 0, \pm 1, \pm 2, \dots$

Se puede observar que:

$$P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1) = \beta_{k_1} \beta_{k_2 - k_1} \tag{1.24}$$

siendo $k_0 = 0$.

En efecto:

Usando el resultado de la probabilidad conjunta de dos v.a. tenemos:

$$P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1) = P(Y_{i,k_1}^t = 1) P(Y_{i,k_2}^t = 1 | Y_{i,k_1}^t = 1) \tag{1.25}$$

ahora, suponiendo que el ancho de intervalo donde ocurre la probabilidad es $k_i - k_{i-1}$:

$$P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1) = \beta_{k_1} \beta_{k_2 - k_1} \blacksquare$$

DEFINICIÓN 1.6 *Asumiendo las definiciones anteriores, el proceso INMA(2) admite la siguiente forma:*

$$X_t = \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2} + \epsilon_t$$

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

$$\beta_j * \epsilon_t = \sum_{i=1}^{\epsilon_t} Y_{i,j}^t$$

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

1.4.7. Propiedades Generales del Modelo de Media Móvil de orden 2, INMA(2).

Sin pérdida de generalidad y asumiendo lo detallado en la definición 1.6, definiremos el proceso INMA(2) así:

$$X_t = \beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2}$$

Suponiendo que este modelo es siempre estacionario y asumiendo que ϵ_t tiene media μ_ϵ y varianza finita σ_ϵ^2 , las propiedades se definen:

1. El Valor Esperado está dado por:

$$E(X_t) = \mu_\epsilon (1 + \beta_1 + \beta_2) \quad (1.26)$$

Prueba:

$$\begin{aligned} E(X_t) &= E(\beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2}) \\ &= E(\beta_0 * \epsilon_t) + E(\beta_1 * \epsilon_{t-1}) + E(\beta_2 * \epsilon_{t-2}) \\ &= \beta_0 E(\epsilon_t) + \beta_1 E(\epsilon_{t-1}) + \beta_2 E(\epsilon_{t-2}) \\ &= \beta_0 \mu_\epsilon + \beta_1 \mu_\epsilon + \beta_2 \mu_\epsilon = \mu_\epsilon (1 + \beta_1 + \beta_2) \quad \blacksquare \end{aligned}$$

2. La Varianza está dado por:

$$V(X_t) = \mu_\epsilon [\beta_1 (1 - \beta_1) + \beta_2 (1 - \beta_2)] + \sigma_\epsilon^2 [1 + \beta_1^2 + \beta_2^2] \quad (1.27)$$

Prueba:

$$\begin{aligned} V(X_t) &= V(\beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2} + \epsilon_t) \\ &= V(\beta_1 * \epsilon_{t-1}) + V(\beta_2 * \epsilon_{t-2}) + V(\epsilon_t) \end{aligned}$$

De igual manera que se obtuvo la varianza de un modelo INAR(1) y suponiendo que $V(\epsilon_t) = V(\epsilon_{t-1}) = V(\epsilon_{t-2}) = \sigma_\epsilon^2$ y $E(\epsilon_{t-1}) = E(\epsilon_{t-2}) = \mu_\epsilon$, se tiene:

$$\begin{aligned} V(\beta_1 * \epsilon_{t-1}) &= \beta_1^2 V(\epsilon_{t-1}) + \beta_1(1 - \beta_1) E(\epsilon_{t-1}) \\ V(\beta_2 * \epsilon_{t-2}) &= \beta_2^2 V(\epsilon_{t-2}) + \beta_2(1 - \beta_2) E(\epsilon_{t-2}) \end{aligned}$$

por lo que la varianza de X_t se expresa:

$$\begin{aligned} V(X_t) &= \beta_1^2 V(\epsilon_{t-1}) + \beta_1(1 - \beta_1) E(\epsilon_{t-1}) + \beta_2^2 V(\epsilon_{t-2}) + \beta_2(1 - \beta_2) E(\epsilon_{t-2}) \\ &+ V(\epsilon_t) \\ &= \beta_1^2 \sigma_\epsilon^2 + \beta_1(1 - \beta_1) \mu_\epsilon + \beta_2^2 \sigma_\epsilon^2 + \beta_2(1 - \beta_2) \mu_\epsilon + \sigma_\epsilon^2 \\ &= \mu_\epsilon [\beta_1(1 - \beta_1) + \beta_2(1 - \beta_2)] + \sigma_\epsilon^2 [1 + \beta_1^2 + \beta_2^2] \quad \blacksquare \end{aligned}$$

3. La Covarianza está dado por:

$$\gamma(k) = \begin{cases} \mu_\epsilon \beta_1 (\beta_1 - \beta_2) + \beta_1 \beta_2 \sigma_\epsilon^2 + \beta_1 \sigma_\epsilon^2 & \text{si } k = 1 \\ \beta_2 \sigma_\epsilon^2 & \text{si } k = 2 \\ 0 & \text{si } k > 2 \end{cases} \quad (1.28)$$

Prueba:

$$\begin{aligned} Cov(X_t, X_{t-k}) &= Cov(\beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2}, \\ &\quad \beta_0 * \epsilon_{t-k} + \beta_1 * \epsilon_{t-k-1} + \beta_2 * \epsilon_{t-k-2}) \end{aligned}$$

aplicando propiedades de la covarianza, tenemos:

$$\begin{aligned}
Cov(X_{t-k}, X_t) &= Cov(\beta_0 * \epsilon_t, \beta_0 * \epsilon_{t-k}) + Cov(\beta_0 * \epsilon_t, \beta_1 * \epsilon_{t-k-1}) \\
&+ Cov(\beta_0 * \epsilon_t, \beta_2 * \epsilon_{t-k-2}) \\
&+ Cov(\beta_1 * \epsilon_{t-1}, \beta_0 * \epsilon_{t-k}) + Cov(\beta_1 * \epsilon_{t-1}, \beta_1 * \epsilon_{t-k-1}) \\
&+ Cov(\beta_1 * \epsilon_{t-1}, \beta_2 * \epsilon_{t-k-2}) \\
&+ Cov(\beta_2 * \epsilon_{t-2}, \beta_0 * \epsilon_{t-k}) + Cov(\beta_2 * \epsilon_{t-2}, \beta_1 * \epsilon_{t-k-1}) \\
&+ Cov(\beta_2 * \epsilon_{t-2}, \beta_2 * \epsilon_{t-k-2})
\end{aligned} \tag{1.29}$$

para $k = 0$ se tiene la varianza (1.27), por lo que nos interesa $k = 1, 2, \dots$ sabemos que ϵ_t son v.a. i.i.d. entonces la covarianza para retardos diferentes de ϵ_t es cero, así:

$$a) Cov(X_{t-1}, X_t) = Cov(\beta_0 * \epsilon_{t-1}, \beta_1 * \epsilon_{t-1}) + Cov(\beta_1 * \epsilon_{t-2}, \beta_2 * \epsilon_{t-2})$$

$$b) Cov(X_{t-2}, X_t) = Cov(\beta_0 * \epsilon_{t-2}, \beta_2 * \epsilon_{t-2})$$

$$c) Cov(X_{t-3}, X_t) = 0$$

Un resultado que obtienen Al-Osh y Alzaid (1988b) para encontrar la covarianza:

$$\begin{aligned}
Cov(\beta_j * \epsilon_t, \beta_{j+k} * \epsilon_t / \epsilon_t = x) &= Cov\left(\sum_{i=1}^x Y_{i,j}, \sum_{i=1}^x Y_{i,j+k}\right) \\
&= Cov(Y_{1,j} + Y_{2,j} + \dots + Y_{x,j}, \\
&\quad Y_{1,j+k} + Y_{2,j+k} + \dots + Y_{x,j+k}) \\
&= Cov(Y_{1,j}, Y_{1,j+k}) + Cov(Y_{2,j}, Y_{2,j+k}) + \dots \\
&+ Cov(Y_{x,j}, Y_{x,j+k}) \\
&= \sum_{i=1}^x Cov(Y_{i,j}, Y_{i,j+k}) = \sum_{i=1}^x [E(Y_{i,j} Y_{i,j+k}) - \beta_j \beta_{j+k}]
\end{aligned}$$

usando la ecuación (1.25):

$$\begin{aligned}
E(Y_{i,j} = 1, Y_{i,j+k} = 1) &= (Y_{i,j} = 1) (Y_{i,j+k} = 1) P(Y_{i,j} = 1, Y_{i,j} = 1) \\
&= P(Y_{i,j} = 1) P(Y_{i,j+k} = 1 / Y_{i,j} = 1) = \beta_j \beta_{j+k-j} = \beta_j \beta_k
\end{aligned}$$

se verifica que:

$$Cov(\beta_j * \epsilon_t, \beta_{j+k} * \epsilon_t / \epsilon_t = x) = \sum_{i=1}^x \beta_j (\beta_k - \beta_{j+k}) = x \beta_j (\beta_k - \beta_{j+k})$$

usando este último resultado, tenemos:

$$Cov(\beta_j * \epsilon_t, \beta_{j+k} * \epsilon_t) = \beta_j (\beta_k - \beta_{j+k}) \mu_\epsilon + \beta_j \beta_{j+k} \sigma_\epsilon^2 \quad (1.30)$$

Para el literal a), se tiene:

$$Cov(\beta_0 * \epsilon_{t-1}, \beta_1 * \epsilon_{t-1}) = \beta_0 (\beta_1 - \beta_1) \mu_\epsilon + \beta_0 \beta_1 \sigma_\epsilon^2 = \beta_1 \sigma_\epsilon^2$$

$$Cov(\beta_1 * \epsilon_{t-2}, \beta_2 * \epsilon_{t-2}) = \beta_1 (\beta_1 - \beta_2) \mu_\epsilon + \beta_1 \beta_2 \sigma_\epsilon^2$$

de estos resultados, se observa que:

$$Cov(X_{t-1}, X_t) = \beta_1 \sigma_\epsilon^2 + \beta_1 (\beta_1 - \beta_2) \mu_\epsilon + \beta_1 \beta_2 \sigma_\epsilon^2 \quad \blacksquare$$

Para el literal b), se tiene:

$$Cov(X_{t-2}, X_t) = Cov(\beta_0 * \epsilon_{t-2}, \beta_2 * \epsilon_{t-2}) = \beta_0 (\beta_2 - \beta_2) \mu_\epsilon + \beta_0 \beta_2 \sigma_\epsilon^2 = \beta_2 \sigma_\epsilon^2 \quad \blacksquare$$

Para el literal c), el resultado se observa de la ecuación (1.29), donde los valores de ϵ_{t-k} no están correlados para $k > 2$. ■

4. La Autocorrelación está dado por:

$$\rho(k) = \begin{cases} \frac{\beta_1 \beta_2 \sigma_\epsilon^2 + \beta_1 (\beta_1 - \beta_2) \mu_\epsilon + \beta_1 \sigma_\epsilon^2}{[\beta_1 (1 - \beta_1) + \beta_2 (1 - \beta_2)] \mu_\epsilon + [1 + \beta_1^2 + \beta_2^2] \sigma_\epsilon^2} & \text{si } k = 1 \\ \frac{\beta_2 \sigma_\epsilon^2}{[\beta_1 (1 - \beta_1) + \beta_2 (1 - \beta_2)] \mu_\epsilon + [1 + \beta_1^2 + \beta_2^2] \sigma_\epsilon^2} & \text{si } k = 2 \\ 0 & \text{si } k > 2 \end{cases} \quad (1.31)$$

Este resultado se obtiene de dividir la respectiva covarianza entre la varianza del modelo INMA(2). ■

1.4.8. Definición del Modelo de Media Móvil de orden q , INMA(q).

Para dar mayor flexibilidad a la modelación de series temporales de valores enteros no negativos, Al-Osh y Alzaid (1988b) presentan una extensión del proceso INMA(1).

Para ello, consideran un proceso en el que la máxima vida para un elemento en el sistema es $(q + 1)$ unidades de tiempo. En primer lugar, se consideran α_j ($j = 1, 2, \dots, q$) 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 ($i = 0, 1, 2, \dots, q$) como:

$$\beta_0 = 1$$

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

Los coeficientes β_i pueden interpretarse como la probabilidad que tiene un elemento de ϵ_t de ser un elemento de X_{t+i} (este elemento podría estar presente en el sistema durante instantes precedentes al $t + i$).

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

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

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

Se puede observar que:

$$P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1, \dots, Y_{i,k_j}^t = 1) = \prod_{l=1}^j \beta_{k_l - k_{l-1}} \quad (1.32)$$

siendo $k_0 = 0$.

En efecto:

generalizando el resultado (1.25), tenemos:

$$\begin{aligned} P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1, \dots, Y_{i,k_j}^t = 1) &= P(Y_{i,k_1}^t = 1) P(Y_{i,k_2}^t = 1 | Y_{i,k_1}^t = 1) \\ &\times P(Y_{i,k_3}^t = 1 | Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1) \dots \\ &\times P(Y_{i,k_j}^t = 1 | Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1, \dots, Y_{i,k_{j-1}}^t = 1) \end{aligned}$$

utilizando la propiedad de los procesos de Markov estacionarios $\{Z_t : t \in T \subseteq (-\infty, \infty)\}$:

$$P(Z_t | Z_{t-1}, Z_{t-2}, \dots) = P(Z_t | Z_{t-1})$$

donde, la probabilidad de un suceso en tiempo t , conociendo la historia pasada del suceso, solo depende únicamente del suceso próximo anterior ($Y_{i,j}^t$, ($i = 1, 2, \dots$) y ($j = 1, 2, \dots, q$) cumple las condiciones del proceso de Markov), entonces:

$$\begin{aligned} P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1, \dots, Y_{i,k_j}^t = 1) &= P(Y_{i,k_1}^t = 1) P(Y_{i,k_2}^t = 1 | Y_{i,k_1}^t = 1) \\ &\times P(Y_{i,k_3}^t = 1 | Y_{i,k_2}^t = 1) P(Y_{i,k_4}^t = 1 | Y_{i,k_3}^t = 1) \dots \\ &\times P(Y_{i,k_j}^t = 1 | Y_{i,k_{j-1}}^t = 1) \end{aligned}$$

ahora, suponiendo que el ancho de intervalo donde ocurre la probabilidad es $k_i - k_{i-1}$:

$$\begin{aligned} P(Y_{i,k_1}^t = 1, Y_{i,k_2}^t = 1, \dots, Y_{i,k_j}^t = 1) &= \beta_{k_1} \beta_{k_2 - k_1} \beta_{k_3 - k_2} \beta_{k_4 - k_3} \dots \beta_{k_j - k_{j-1}} \\ &= \prod_{l=1}^j \beta_{k_l - k_{l-1}} \quad \blacksquare \end{aligned}$$

DEFINICIÓN 1.7 *Asumiendo las definiciones anteriores, el proceso INMA(q) admite la siguiente forma:*

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

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

$$\beta_j * \epsilon_t = \sum_{i=1}^{\epsilon_t} Y_{i,j}^t$$

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

1.4.9. Propiedades Generales del Modelo de Media Móvil de orden q , INMA(q).

Sin pérdida de generalidad y asumiendo lo detallado en la definición (1.7), definiremos el proceso INMA(q) así:

$$X_t = \beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2} + \dots + \beta_q * \epsilon_{t-q}$$

Suponiendo que este modelo es siempre estacionario y que ϵ_t tiene media μ_ϵ y varianza finita σ_ϵ^2 , las propiedades se definen:

1. El Valor Esperado está dado por:

$$E(X_t) = \mu_\epsilon \sum_{j=0}^q \beta_j \quad (1.33)$$

Prueba:

$$\begin{aligned} E(X_t) &= E(\beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \dots + \beta_q * \epsilon_{t-q}) \\ &= E(\beta_0 * \epsilon_t) + E(\beta_1 * \epsilon_{t-1}) + \dots + E(\beta_q * \epsilon_{t-q}) \\ &= \beta_0 E(\epsilon_t) + \beta_1 E(\epsilon_{t-1}) + \dots + \beta_q E(\epsilon_{t-q}) \\ &= \beta_0 \mu_\epsilon + \beta_1 \mu_\epsilon + \dots + \beta_q \mu_\epsilon = \mu_\epsilon \sum_{j=0}^q \beta_j \quad \blacksquare \end{aligned}$$

2. La Varianza está dado por:

$$V(X_t) = \mu_\epsilon \sum_{j=0}^q \beta_j (1 - \beta_j) + \sigma_\epsilon^2 \sum_{j=0}^q \beta_j^2 \quad (1.34)$$

Prueba:

$$\begin{aligned} V(X_t) &= V(\beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \dots + \beta_q * \epsilon_{t-q}) \\ &= V(\beta_0 * \epsilon_t) + V(\beta_1 * \epsilon_{t-1}) + \dots + V(\beta_q * \epsilon_{t-q}) \end{aligned}$$

Como hipótesis tenemos que: $V(\epsilon_t) = V(\epsilon_{t-1}) = \dots = V(\epsilon_{t-q}) = \sigma_\epsilon^2$ y $E(\epsilon_t) = E(\epsilon_{t-1}) = \dots = E(\epsilon_{t-q}) = \mu_\epsilon$ y recordamos:

$$V(\beta_i * \epsilon_{t-i}) = \beta_i^2 V(\epsilon_{t-i}) + \beta_i (1 - \beta_i) E(\epsilon_{t-i}) \quad \text{para } i = 0, 1, \dots, q$$

por lo que la varianza se expresa:

$$\begin{aligned} V(X_t) &= \beta_0^2 V(\epsilon_t) + \beta_0 (1 - \beta_0) E(\epsilon_t) + \beta_1^2 V(\epsilon_{t-1}) + \beta_1 (1 - \beta_1) E(\epsilon_{t-1}) \\ &\quad + \dots + \beta_q^2 V(\epsilon_{t-q}) + \beta_q (1 - \beta_q) E(\epsilon_{t-q}) \\ &= \beta_0^2 \sigma_\epsilon^2 + \beta_0 (1 - \beta_0) \mu_\epsilon + \beta_1^2 \sigma_\epsilon^2 + \beta_1 (1 - \beta_1) \mu_\epsilon + \dots + \\ &\quad \beta_q^2 \sigma_\epsilon^2 + \beta_q (1 - \beta_q) \mu_\epsilon \\ &= \mu_\epsilon [\beta_0 (1 - \beta_0) + \beta_1 (1 - \beta_1) + \dots + \beta_q (1 - \beta_q)] + \\ &\quad \sigma_\epsilon^2 [\beta_0^2 + \beta_1^2 + \dots + \beta_q^2] \quad \blacksquare \end{aligned}$$

3. La Covarianza está dado por:

$$\gamma(k) = \begin{cases} \sum_{j=0}^{q-k} [\beta_j (\beta_k - \beta_{k+j}) \mu_\epsilon + \beta_j \beta_{k+j} \sigma_\epsilon^2] & \text{si } k = 1, 2, \dots, q \\ 0 & \text{si } k > q \end{cases} \quad (1.35)$$

Obsérvese que X_t y X_{t-k} están dados por:

$$\begin{aligned} X_t &= \beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \dots + \beta_{q-1} * \epsilon_{t-q+1} + \beta_q * \epsilon_{t-q} \\ X_{t-k} &= \beta_0 * \epsilon_{t-k} + \beta_1 * \epsilon_{t-k-1} + \dots + \beta_{q-1} * \epsilon_{t-k-(q-1)} + \beta_q * \epsilon_{t-k-q} \end{aligned}$$

sustituyendo se tiene:

$$\begin{aligned} \gamma(k) = Cov(X_t, X_{t-k}) &= Cov(\beta_0 * \epsilon_t + \beta_1 * \epsilon_{t-1} + \dots + \beta_{q-1} * \epsilon_{t-q+1} + \beta_q * \epsilon_{t-q}, \\ &\quad \beta_0 * \epsilon_{t-k} + \beta_1 * \epsilon_{t-k-1} + \dots + \beta_{q-1} * \epsilon_{t-k-(q-1)} + \beta_q * \epsilon_{t-k-q}) \end{aligned}$$

podemos notar que dependiendo del valor de k , se obtendrá el número de covarianzas, de manera que si $k = 1$, se obtiene la suma de (q) covarianzas, y se observa que para el proceso X_t no influye el primer sumando $\beta_0 * \epsilon_t$ y en el proceso X_{t-k} no influye el último sumando $\beta_q * \epsilon_{t-k-q}$, es decir:

$$\begin{aligned} \gamma(1) = Cov(X_t, X_{t-1}) &= Cov(\beta_1 * \epsilon_{t-1}, \beta_0 * \epsilon_{t-1}) + Cov(\beta_2 * \epsilon_{t-2}, \beta_1 * \epsilon_{t-2}) + \dots \\ &+ Cov(\beta_q * \epsilon_{t-q}, \beta_{q-1} * \epsilon_{t-q}) \end{aligned}$$

ahora, si $k = 2$, se obtiene la suma de $(q - 1)$ covarianzas, y se observa que para el proceso X_t no influye el primer y segundo sumando $\beta_0 * \epsilon_t$ y $\beta_1 * \epsilon_{t-1}$ y en el proceso X_{t-k} no influye el penúltimo y último sumando $\beta_{q-1} * \epsilon_{t-k-(q-1)}$ y $\beta_q * \epsilon_{t-k-q}$, es decir:

$$\begin{aligned} \gamma(2) = Cov(X_t, X_{t-2}) &= Cov(\beta_2 * \epsilon_{t-2}, \beta_0 * \epsilon_{t-2}) + Cov(\beta_3 * \epsilon_{t-3}, \beta_1 * \epsilon_{t-3}) + \dots \\ &+ Cov(\beta_q * \epsilon_{t-q}, \beta_{q-1} * \epsilon_{t-q}) \end{aligned}$$

así sucesivamente, se obtienen las covarianzas de retardo k , se observa que a medida aumenta el orden del retardo, disminuye la cantidad de sumandos, de aquí verificamos

el orden del sumatorio en la expresión (1.35).

Por lo tanto:

$$\gamma(k) = Cov(X_t, X_{t-k}) = \sum_{j=0}^{q-k} Cov(\beta_j * \epsilon_{t-j}, \beta_{j+k} * \epsilon_{t-j})$$

Dado que el proceso INMA(q) es siempre estacionario en media, la expresión anterior se puede escribir por comodidad, así:

$$\gamma(k) = Cov(X_t, X_{t-k}) = \sum_{j=0}^{q-k} Cov(\beta_j * \epsilon_t, \beta_{j+k} * \epsilon_t)$$

Generalizando el resultado de la ecuación (1.30) que obtienen Al-Osh y Alzaid (1988b) para encontrar la covarianza de un modelo INMA(2) y usando la expresión (1.32), tenemos:

$$Cov(\beta_j * \epsilon_t, \beta_{j+k} * \epsilon_t) = \beta_j (\beta_k - \beta_{k+j}) \mu_\epsilon + \beta_j \beta_{k+j} \sigma_\epsilon^2$$

usando este resultado, la covarianza de retardo k para un proceso INMA(q) donde $k = 1, 2, \dots, q$ es:

$$\gamma(k) = Cov(X_t, X_{t-k}) = \sum_{j=0}^{q-k} [\beta_j (\beta_k - \beta_{k+j}) \mu_\epsilon + \beta_j \beta_{k+j} \sigma_\epsilon^2] \quad \blacksquare$$

4. La Autocorrelación está dado por:

$$\rho(k) = \begin{cases} 1 & \text{si } k = 0 \\ \frac{\sum_{j=0}^{q-k} [\beta_j (\beta_k - \beta_{j+k}) \mu_\epsilon + \beta_j \beta_{j+k} \sigma_\epsilon^2]}{\mu_\epsilon \sum_{j=0}^q \beta_j (1 - \beta_j) + \sigma_\epsilon^2 \sum_{j=0}^q \beta_j^2} & \text{si } 1 \leq k \leq q. \\ 0 & \text{si } k > q \end{cases} \quad (1.36)$$

Este resultado se obtiene de dividir la respectiva covarianza entre la varianza del modelo INMA(q). \blacksquare

La función de autocorrelación del proceso INMA(q) es similar al MA(q) estándar, es decir, esta función se anulan cuando el defase temporal considerado es superior al orden del modelo.

1.5. Modelo mixto Autorregresivo y de Media Móvil de Valores Enteros de orden 1. INARMA(1,1).

1.5.1. Definición del Modelo mixto Autorregresivo y de Media Móvil de Valores Enteros de orden 1. INARMA(1,1).

Discutiremos un modelo que se construye uniendo dos procesos, el INAR(1) y el INMA(1) definidos en los apartados anteriores y discutidos por Al-Osh y Alzaid (1988b) considerando un proceso de innovación común, $\{\epsilon_t, t \in \mathbb{Z}\}$.

DEFINICIÓN 1.8 Sea $\{X_t; t = \pm 1, \pm 2, \pm 3, \dots\}$ una sucesión de v.a. de valores enteros no negativos, entonces el proceso INARMA(1,1) admite la forma:

$$X_t = Z_{t-1} + \beta * \epsilon_t$$

donde:

$$Z_{t-1} = \alpha * Z_{t-2} + \epsilon_{t-1}$$

donde ϵ_t es v.a. i.i.d. de media μ_ϵ y varianza σ_ϵ^2 .

Notemos que si $\alpha \in [0, 1)$, $\mu_\epsilon < \infty$ y $\sigma_\epsilon^2 < \infty$ el proceso $\{Z_t\}$ es estacionario, además, el proceso INMA(1), definido en el apartado 1.2 es siempre estacionario, por lo tanto $\{X_t, t \in \mathbb{Z}\}$ también será estacionario.

Esta forma de representar el modelo INARMA(1,1) evita contar más de una vez un elemento en el sistema, ya que si definimos el modelo INARMA(1,1) de forma análoga al modelo ARMA(1,1) estándar, tenemos:

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

Observamos que un elemento de ϵ_{t-1} puede entrar al sistema en el tiempo t a través de $\alpha * X_{t-1}$ ó $\beta * \epsilon_{t-1}$, por lo que cabe la posibilidad de contar un elemento dos veces. Así la definición (1.8) excluye esa posibilidad.

1.5.2. Propiedades Generales del Modelo INARMA(1,1).

Considerando las propiedades de los Modelos INAR(1) e INMA(1), haciendo uso de la independencia de las v.a. $\beta * \epsilon_{t-i}$ y $\beta * \epsilon_{t-j}$ que son v.a. independientes para $i \neq j$ y $E(\epsilon_t) = \mu_\epsilon < \infty$, $V(\epsilon_t) = \sigma_\epsilon^2 < \infty$. Las propiedades del modelo mixto autorregresivo y de media móvil de orden 1, INARMA(1,1), se definen:

1. El Valor Esperado está dado por:

$$E(X_t) = \frac{\mu_\epsilon}{1 - \alpha} + \beta\mu_\epsilon \quad (1.37)$$

Prueba:

$$E(X_t) = E(Z_{t-1} + \beta * \epsilon_t) = E(Z_{t-1}) + E(\beta * \epsilon_t) = E(Z_{t-1}) + \beta E(\epsilon_t)$$

sustituyendo el valor esperado de un modelo INAR(1) y $E(\epsilon_t) = \mu_\epsilon$, tenemos:

$$E(X_t) = E(Z_{t-1}) + \beta E(\epsilon_t) = \frac{\mu_\epsilon}{1 - \alpha} + \beta\mu_\epsilon \quad \blacksquare$$

2. La Varianza está dado por:

$$V(X_t) = \frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} + [\beta(1 - \beta)\mu_\epsilon + \beta^2\sigma_\epsilon^2] \quad (1.38)$$

Prueba:

$$V(X_t) = V(Z_{t-1} + \beta * \epsilon_t) = V(Z_{t-1}) + V(\beta * \epsilon_t)$$

aplicando la varianza condicional, de igual manera que para un modelo INMA(1), tenemos:

$$V(\beta * \epsilon_t) = \beta^2 V(\epsilon_{t-1}) + \beta(1 - \beta) E(\epsilon_{t-1})$$

sustituyendo la varianza del modelo INAR(1) y la expresión anterior:

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

sabiendo que $E(\epsilon_t) = \mu_\epsilon$ y $V(\epsilon_t) = \sigma_\epsilon^2$:

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

3. La Covarianza está dado por:

$$\gamma(k) = \begin{cases} \left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha + \beta\sigma_\epsilon^2 & \text{si } k = 1 \\ \left(\left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha + \beta\sigma_\epsilon^2 \right) \alpha^{k-1} & \text{si } k \geq 2 \end{cases} \quad (1.39)$$

Prueba:

Primero probaremos el resultado para $k = 1$:

expresando X_t y X_{t-1} como la definición 1.8:

$$X_t = Z_{t-1} + \beta * \epsilon_t \quad \text{con } Z_{t-1} = \alpha * Z_{t-2} + \epsilon_{t-1}$$

$$X_{t-1} = Z_{t-2} + \beta * \epsilon_{t-1} \quad \text{con } Z_{t-2} = \alpha * Z_{t-3} + \epsilon_{t-2}$$

sustituyendo Z_{t-1} en X_t y aplicando la propiedad de la covarianza:

$$\begin{aligned} \gamma(1) &= Cov(Z_{t-1} + \beta * \epsilon_t, Z_{t-2} + \beta * \epsilon_{t-1}) \\ &= Cov(\alpha * Z_{t-2} + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-2} + \beta * \epsilon_{t-1}) \\ &= Cov(\alpha * Z_{t-2}, Z_{t-2}) + Cov(\epsilon_{t-1}, \beta * \epsilon_{t-1}) \end{aligned}$$

expresando la covarianza en términos de valor esperado, se obtiene la varianza de un modelo INAR(1) e INMA(1), respectivamente:

$$\gamma(1) = Cov(\alpha * Z_{t-2}, Z_{t-2}) + Cov(\epsilon_{t-1}, \beta * \epsilon_{t-1}) = \left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha + \beta\sigma_\epsilon^2 \quad \blacksquare$$

Ahora probaremos el resultado para $k \geq 2$:

expresando X_t , X_{t-1} y X_{t-k} como la definición (1.8):

$$\begin{aligned} X_t &= Z_{t-1} + \beta * \epsilon_t \quad \text{con } Z_{t-1} = \alpha * Z_{t-2} + \epsilon_{t-1} \\ X_{t-1} &= Z_{t-2} + \beta * \epsilon_{t-1} \quad \text{con } Z_{t-2} = \alpha * Z_{t-3} + \epsilon_{t-2} \\ &\vdots \\ X_{t-k} &= Z_{t-k-1} + \beta * \epsilon_{t-k} \quad \text{con } Z_{t-k-1} = \alpha * Z_{t-k-2} + \epsilon_{t-k-1} \end{aligned}$$

sustituyendo Z_{t-1} , $Z_{t-2} \dots Z_{t-k}$ en X_t y aplicando la propiedad (6) del operador de

refinamiento:

$$\begin{aligned}
\gamma(k) &= Cov(\alpha * Z_{t-2} + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha * (\alpha * Z_{t-3} + \epsilon_{t-2}) + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha^2 * Z_{t-3} + \alpha * \epsilon_{t-2} + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha^2 * (\alpha * Z_{t-4} + \epsilon_{t-3}) + \alpha * \epsilon_{t-2} + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha^3 * Z_{t-4} + \alpha^2 * \epsilon_{t-3} + \alpha * \epsilon_{t-2} + \epsilon_{t-1} + \beta * \epsilon_t, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\vdots \\
&\stackrel{d}{=} Cov(\alpha^k * Z_{t-k-1} + \alpha^{k-1} * \epsilon_{t-k} + \alpha^{k-2} * \epsilon_{t-k+1} + \dots + \epsilon_{t-1} + \beta * \epsilon_t, \\
&\quad Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha^k * Z_{t-k-1} + \alpha^{k-1} * \epsilon_{t-k}, Z_{t-k-1} + \beta * \epsilon_{t-k}) \\
&\stackrel{d}{=} Cov(\alpha^k * Z_{t-k-1}, Z_{t-k-1}) + Cov(\alpha^{k-1} * \epsilon_{t-k}, \beta * \epsilon_{t-k})
\end{aligned}$$

de la ecuación (1.4) de un modelo INAR(1), se tiene:

$$Cov(\alpha^k * Z_{t-k-1}, Z_{t-k-1}) = \left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha^k$$

por otra parte:

$$Cov(\alpha^{k-1} * \epsilon_{t-k}, \beta * \epsilon_{t-k}) = E[(\alpha^{k-1} * \epsilon_{t-k})(\beta * \epsilon_{t-k})] - \alpha^{k-1}\beta(E(\epsilon_{t-k}))^2$$

usando la propiedad (5), del operador de refinamiento, tenemos:

$$E[(\alpha^{k-1} * \epsilon_{t-k})(\beta * \epsilon_{t-k})] = \alpha^{k-1}\beta E[\epsilon_{t-k}^2]$$

entonces:

$$\begin{aligned}
Cov(\alpha^{k-1} * \epsilon_{t-k}, \beta * \epsilon_{t-k}) &= \alpha^{k-1}\beta E[\epsilon_{t-k}^2] - \alpha^{k-1}\beta(E(\epsilon_{t-k}))^2 \\
&= \alpha^{k-1}\beta V(\epsilon_{t-k}) = \alpha^{k-1}\beta\sigma_\epsilon^2
\end{aligned}$$

finalmente, se tiene:

$$\gamma(k) \stackrel{d}{=} \left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha^k + \alpha^{k-1}\beta\sigma_\epsilon^2 \stackrel{d}{=} \left(\left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1 - \alpha^2} \right) \alpha + \beta\sigma_\epsilon^2 \right) \alpha^{k-1} \quad \blacksquare$$

4. La Autocorrelación está dado por:

$$\rho(k) = \begin{cases} \frac{\left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2}\right) \alpha + \beta\sigma_\epsilon^2}{\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2} + [\beta(1-\beta)\mu_\epsilon + (1+\beta^2)\sigma_\epsilon^2]} & \text{si } k = 1 \\ \frac{\left(\left(\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2}\right) \alpha + \beta\sigma_\epsilon^2\right) \alpha^{k-1}}{\frac{\alpha\mu_\epsilon + \sigma_\epsilon^2}{1-\alpha^2} + [\beta(1-\beta)\mu_\epsilon + (1+\beta^2)\sigma_\epsilon^2]} & \text{si } k \geq 2 \end{cases} \quad (1.40)$$

En efecto:

Este resultado se obtiene de dividir la ecuación (1.39) entre la ecuación (1.38). ■

Capítulo 2

Estimación de Parámetros de los Modelos INARMA.

2.1. Estimación de Parámetros del Modelo INAR(1) con marginal de Poisson.

Recordemos que el modelo INAR(1) está dado por:

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

La Función de Autocorrelación está dada por:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \alpha^k$$

2.1.1. Estimadores de Yule-Walker.

La manera más sencilla de obtener un estimador para α , es reemplazando $\hat{\alpha}$ en $\rho(k)$, reemplazando α en la ecuación de la covarianza muestral de primer orden ($\alpha = \rho(1)$), notamos que la función de autocovarianza muestral es parecida a la ecuación de Yule-Walker para

datos continuos, por lo que obtenemos:

$$\hat{\alpha} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-1} (X_t - \bar{X}_t) (X_{t+1} - \bar{X}_t)}{\sum_{t=1}^n (X_t - \bar{X}_t)^2} \quad (2.1)$$

donde \bar{X}_t es la media muestral.

Ahora el estimador para λ se puede obtener:

$$\hat{\lambda} = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha} X_{t-1}) \quad (2.2)$$

En efecto:

$$\begin{aligned} \epsilon_t &= X_t - \alpha * X_{t-1} \\ E(\epsilon_t) &= E(X_t) - \alpha E(X_{t-1}) \end{aligned}$$

ahora, un estimador razonable para λ es:

$$\begin{aligned} E(\widehat{\epsilon}_t) &= \sum_{t=1}^n \frac{X_t}{n} - \hat{\alpha} \sum_{t=1}^n \frac{X_{t-1}}{n} \\ &= \frac{1}{n} \left(\sum_{t=1}^n X_t - \hat{\alpha} \sum_{t=1}^n X_{t-1} \right) \\ &= \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha} X_{t-1}) \end{aligned}$$

ya que ϵ_t tiene una distribución de Poisson de parámetro λ . ■

2.1.2. Estimadores Mínimos Cuadrados Condicionales.

El método de mínimos cuadrados condicionales fue desarrollado por Klimko y Nelson (1978) y se basa en la minimización de la suma de desviaciones cuadradas acerca de la esperanza condicional.

Para encontrar los estimadores de α y λ de un modelo INAR(1), mediante el método de mínimos cuadrados condicionales, hacemos uso de la media condicional de X_t , dado X_{t-1} la cual está dada por:

$$E[X_t/X_{t-1}] = \alpha X_{t-1} + \lambda \tag{2.3}$$

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

$$Q_n(\alpha, \lambda) = \sum_{t=1}^n (X_t - E[X_t/X_{t-1}])^2 \tag{2.4}$$

con respecto a α y λ .

Derivando (2.4) con respecto a α y λ se tiene:

$$\hat{\alpha} = \frac{n \sum_{t=1}^n X_t X_{t-1} - \sum_{t=1}^n X_t \sum_{t=1}^n X_{t-1}}{n \sum_{t=1}^n X_{t-1}^2 - \left(\sum_{t=1}^n X_{t-1} \right)^2} \tag{2.5}$$

y

$$\hat{\lambda} = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha} X_{t-1}) \tag{2.6}$$

Verificación:

Usando la ecuación (2.4), se tiene:

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

derivando $Q_n(\alpha, \lambda)$ con respecto a α y λ , tenemos:

$$\frac{\partial Q_n(\alpha, \lambda)}{\partial \alpha} = \sum_{t=1}^n (-2X_t X_{t-1} + 2\alpha X_{t-1}^2 + 2\lambda X_{t-1}) = 0 \quad (2.7)$$

$$\frac{\partial Q_n(\alpha, \lambda)}{\partial \lambda} = \sum_{t=1}^n (-2X_t + 2\alpha X_{t-1} + 2\lambda) = 0 \quad (2.8)$$

de la derivada de $Q_n(\alpha, \lambda)$ con respecto a λ se tiene que:

$$\hat{\lambda} = \frac{1}{n} \left[\sum_{t=1}^n X_t - \hat{\alpha} \sum_{t=1}^n X_{t-1} \right] \quad \blacksquare$$

sustituyendo el estimador de $\hat{\lambda}$ en la ecuación (2.7) tenemos para $\hat{\alpha}$:

$$\begin{aligned} \frac{\partial Q_n(\alpha, \lambda)}{\partial \alpha} &= -2 \sum_{t=1}^n X_t X_{t-1} + 2\alpha \sum_{t=1}^n X_{t-1}^2 + 2\lambda \sum_{t=1}^n X_{t-1} = 0 \\ \Rightarrow & - \sum_{t=1}^n X_t X_{t-1} + \hat{\alpha} \sum_{t=1}^n X_{t-1}^2 + \frac{1}{n} \left(\sum_{t=1}^n X_t - \hat{\alpha} \sum_{t=1}^n X_{t-1} \right) \sum_{t=1}^n X_{t-1} = 0 \\ \hat{\alpha} &= \frac{n \sum_{t=1}^n X_t X_{t-1} - \sum_{t=1}^n X_t \sum_{t=1}^n X_{t-1}}{n \sum_{t=1}^n X_{t-1}^2 - \left(\sum_{t=1}^n X_{t-1} \right)^2} \quad \blacksquare \end{aligned}$$

Se esperaría que para procesos estacionarios, las estimaciones de la media y la varianza del modelo sean muy próximos para las primeras n y $n + 1$ observaciones calculadas, vemos que CLS (Estimadores Mínimos Cuadrados Condicionales) en (2.5) y (2.6) son muy parecidos a los estimadores encontrados mediante el proceso de Yule-Walker en (2.1) y (2.2).

2.1.3. Estimadores de Máxima Verosimilitud.

La función de probabilidad (verosimilitud) de una muestra de $(n + 1)$ observaciones $\{X_0, X_1, \dots, X_t\}$ para el proceso INAR(1) puede escribirse como:

$$L(x, \alpha, \lambda) = \left(\prod_{t=1}^n P_t(x_t) \right) \frac{[\lambda/(1 - \alpha)]^{x_0}}{x_0!} e^{-[\lambda/(1-\alpha)]} \quad (2.9)$$

En efecto:

$$L(x, \alpha, \lambda) = P_t(x_t)P_{t-1}(x_{t-1}) \dots P_1(x_1)P_0(x_0)$$

el producto de las funciones de probabilidad de los x_i para $i = 0, 1, 2, \dots, t$.

Del modelo INAR(1) se tiene: $x_0 = \epsilon_0$, dado que solo conocemos los sucesos desde 0 hasta t , asumiremos que $x_{-1} = 0$. Al inicio del apartado, se asumió que ϵ_t tiene una distribución de Poisson y se verificó que el valor esperado del proceso inicial para un modelo estacionario (prueba realizada en el Capítulo I) es $E(x_0) = \mu_\epsilon/(1 - \alpha)$. Entonces, ϵ_0 tendrá una distribución de Poisson con media igual al proceso inicial, por lo que conociendo el valor esperado de x_0 , se sustituye en su distribución:

$$P_0(x_0) = \frac{[\lambda/(1 - \alpha)]^{x_0}}{x_0!} e^{-[\lambda/(1-\alpha)]}$$

este resultado, lo sustituimos en $L(x, \alpha, \lambda)$ y se tiene:

$$L(x, \alpha, \lambda) = \left(\prod_{t=1}^n P_t(x_t) \right) P_0(x_0) = \left(\prod_{t=1}^n P_t(x_t) \right) \frac{[\lambda/(1 - \alpha)]^{x_0}}{x_0!} e^{-[\lambda/(1-\alpha)]} \quad \blacksquare$$

Se puede verificar que:

$$P_t(X_t = x_t) = e^{-\lambda} \sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t - i)!} \binom{x_{t-1}}{i} \alpha^i (1 - \alpha)^{x_{t-1}-i} \quad (2.10)$$

para $t = 1, 2, \dots, n$ y $M = \min(x_{t-1}, x_t)$.

Verificación:

Para verificar este resultado, utilizaremos la distribución de masa de dos variables.

Sea x_1 y x_2 dos variables aleatorias, entonces $x_1 + x_2$ es también variable aleatoria. Si $x_1 + x_2$ es una variable aleatoria discreta, su función de masa de probabilidad viene dada por:

$$P_{x_1+x_2}(y) = \sum_K P_{x_1}(i)P_{x_2}(y - i)$$

donde $y = x_1 + x_2$ y $K = \{y - i \geq 0 / i = 1, 2, \dots\}$ con i la i -ésima observación del conjunto k . Esta función de masa de probabilidad está definida para todo K donde se cumpla:

- $P_{x_1}(i) > 0$
- $P_{x_2}(y - i) > 0$

sustituyendo en $P_t(x_t)$ el modelo INAR(1), se tiene:

$$P_{\alpha * x_{t-1} + \epsilon_t}(x_t) = \sum_{i=0}^M P_{\alpha * x_{t-1}}(i) P_{\epsilon_t}(x_t - i)$$

como sabemos, ϵ_t tiene distribución de Poisson de parámetro $\lambda > 0$, que está definido para todo $i \geq 0$:

$$P_{\epsilon_t}(x_t - i) = \frac{\lambda^{x_t - i}}{(x_t - i)!} e^{-\lambda} \quad (2.11)$$

además, sabemos que $(\alpha * x_{t-1} / x_{t-1})$ tiene distribución binomial con parámetros α y x_{t-1} , entonces si $\epsilon_t = x_t - i$, se tiene que $\alpha * x_{t-1} = i$, por lo tanto:

$$P_{\alpha * x_{t-1}}(i) = \binom{x_{t-1}}{i} \alpha^i (1 - \alpha)^{x_{t-1} - i} \quad (2.12)$$

sustituyendo (2.11) y (2.12) en (2.10), tenemos:

$$P_t(x_t) = e^{-\lambda} \sum_{i=0}^M \frac{\lambda^{x_t - i}}{(x_t - i)!} \binom{x_{t-1}}{i} \alpha^i (1 - \alpha)^{x_{t-1} - i} \quad (2.13)$$

vemos que la sumatoria de esta expresión está definida desde $i = 0$ hasta $M = \min(x_{t-1}, x_t)$, esto se verifica haciendo uso de dos posibles resultados:

- Supongamos que $x_{t-1} < x_t$, tenemos que en la ecuación (2.10) los índices de la sumatoria se definen: $i = 0, 1, 2, \dots$. Vemos que el combinatorio $\binom{x_{t-1}}{i}$ y $(1 - \alpha)^{x_{t-1} - i}$, están definidos para $i = 0, 1, 2, \dots, x_{t-1}$ (no se considera la observación x_t).
- Supongamos que $x_t < x_{t-1}$, tenemos que en la ecuación (2.10) los índices de la sumatoria se definen: $i = 0, 1, 2, \dots$. Vemos que el factorial $(x_t - i)!$ y $\lambda^{x_t - i}$, están definidos para $i = 0, 1, 2, \dots, x_t$ (no se considera la observación x_{t-1}).

De estos resultados se comprueban los índices del sumatorio en la ecuación (2.10).

Ahora, para α y λ dado x_0 , consideraremos los primeros estimadores condicionales de máxima verosimilitud, en este caso la función condicional de verosimilitud se define:

$$L(x, \alpha, \lambda/x_0) = \prod_{t=1}^n P_t(x_t)$$

Sportt (1983) describió un método para maximizar $L(x, \alpha, \lambda/x_0)$ con respecto a α y λ cuando $P_t(x)$ es independiente de t , esto es, el parámetro índice de las observaciones.

Uno de los siguientes resultados obtenidos por Sportt, referente a las derivadas de los logaritmos de las estimaciones de la función condicional de verosimilitud para α y λ son:

$$S_\lambda = \sum_{t=1}^n \frac{\partial \ln P_t(x_t)}{\partial \lambda} = \sum_{t=1}^n \frac{P_t(x_t - 1)}{P_t(x_t)} - n = \sum_{t=1}^n H(t) - n \quad (2.14)$$

$$S_\alpha = \sum_{t=1}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha} = \sum_{t=1}^n \frac{[(x_t - \alpha x_{t-1}) - \lambda H(t)]}{\alpha(1 - \alpha)} \quad (2.15)$$

donde:

$$H(t) = \frac{P_t(x_t - 1)}{P_t(x_t)}$$

Verificación:

Recordemos que $P_t(x_t)$, por la ecuación (2.10) está dada por:

$$P_t(x_t) = e^{-\lambda} \sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i}$$

Ahora, aplicando logaritmo natural a $P_t(x_t)$ se tiene:

$$\ln P_t(x_t) = -\lambda + \ln \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]$$

derivando parcialmente ($\ln P_t(x_t)$) con respecto a λ , sabiendo que $n! = n(n-1)!$, multiplicando y dividiendo la expresión por $e^{-\lambda}$, se tiene:

$$\frac{\partial \ln P_t(x_t)}{\partial \lambda} = -1 + \frac{e^{-\lambda} \left[\sum_{i=0}^M \frac{(x_t-i)\lambda^{x_t-i-1}}{(x_t-i)(x_t-i-1)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]}{e^{-\lambda} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]}$$

de ésta expresión resulta fácil verificar que:

$$S_\lambda = \sum_{t=1}^n \frac{\partial \ln P_t(x_t)}{\partial \lambda} = \sum_{t=1}^n \left[\frac{P_t(x_t - 1)}{P_t(x_t)} \right] - n \quad \blacksquare$$

Para verificar el resultado de S_α , tenemos:

de igual manera que para S_λ aplicamos logaritmo natural en la ecuación (2.10):

$$\ln P_t(x_t) = -\lambda + \ln \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]$$

derivando parcialmente ($\ln P_t(x_t)$) con respecto a α :

$$\frac{\partial \ln P_t(x_t)}{\partial \alpha} = \frac{\left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \left[i\alpha^{i-1}(1-\alpha)^{x_{t-1}-i} - \alpha^i(x_{t-1}-i)(1-\alpha)^{x_{t-1}-i-1} \right] \right]}{\left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]}$$

obteniendo el factor común ($\alpha^i(1-\alpha)^{x_{t-1}-i}$):

$$\frac{\partial \ln P_t(x_t)}{\partial \alpha} = \frac{\left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} (\alpha^i(1-\alpha)^{x_{t-1}-i}) \left[i\alpha^{-1} - (x_{t-1}-i)(1-\alpha)^{-1} \right] \right]}{\left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]}$$

si multiplicamos y dividimos por $e^{-\lambda}$, obtenemos en el denominador $P_t(x_t)$:

$$\begin{aligned} \frac{\partial \ln P_t(x_t)}{\partial \alpha} &= \frac{e^{-\lambda}}{P_t(x_t)} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] \left[\frac{i}{\alpha} - \frac{(x_{t-1}-i)}{(1-\alpha)} \right] \\ &= \frac{e^{-\lambda}}{P_t(x_t)} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] \left[\frac{i - \alpha x_{t-1}}{\alpha(1-\alpha)} \right] \\ &= \frac{e^{-\lambda}}{P_t(x_t)} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] \left[\frac{i - \alpha x_{t-1} + (x_t - x_t)}{\alpha(1-\alpha)} \right] \\ &= \frac{e^{-\lambda}}{P_t(x_t)} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] \left[\frac{(x_t - \alpha x_{t-1}) - (x_t - i)}{\alpha(1-\alpha)} \right] \end{aligned}$$

luego de unos pasos algebraicos, se tiene:

$$= \frac{1}{\alpha(1-\alpha)} \left[(x_t - \alpha x_{t-1}) - \frac{e^{-\lambda} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] (x_t - i)}{P_t(x_t)} \right]$$

la expresión no se altera si hacemos:

$$= \frac{1}{\alpha(1-\alpha)} \left[(x_t - \alpha x_{t-1}) - \frac{\lambda \lambda^{-1} e^{-\lambda} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i}}{(x_t-i)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right] (x_t - i)}{P_t(x_t)} \right]$$

$$= \frac{1}{\alpha(1-\alpha)} \left[(x_t - \alpha x_{t-1}) - \frac{\lambda e^{-\lambda} \left[\sum_{i=0}^M \frac{\lambda^{x_t-i-1}}{(x_t-i-1)!} \binom{x_{t-1}}{i} \alpha^i (1-\alpha)^{x_{t-1}-i} \right]}{P_t(x_t)} \right]$$

por lo que obtenemos $P_t(x_t - 1)$ y $P_t(x_t)$ al igual que para S_λ :

$$\frac{\partial \ln P_t(x_t)}{\partial \alpha} = \frac{1}{\alpha(1-\alpha)} \left[(x_t - \alpha x_{t-1}) - \frac{\lambda P_t(x_t - 1)}{P_t(x_t)} \right]$$

de esta expresión, si $H(t) = \frac{P_t(x_t - 1)}{P_t(x_t)}$ resulta fácil verificar que:

$$S_\alpha = \sum_{t=1}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha} = \sum_{t=1}^n \frac{[(x_t - \alpha x_{t-1}) - \lambda H(t)]}{\alpha(1-\alpha)} \quad \blacksquare$$

Para obtener los estimadores de máxima verosimilitud, hacemos $S_\lambda = 0$ y $S_\alpha = 0$:

$$S_\lambda = \sum_{t=1}^n \left[\frac{P_t(x_t - 1)}{P_t(x_t)} \right] - n = \sum_{t=1}^n H(t) - n \Rightarrow \sum_{t=1}^n H(t) = n$$

sustituyendo n en S_α , se obtiene el estimador de λ :

$$S_{\hat{\alpha}} = \sum_{t=1}^n \left[\frac{(x_t - \hat{\alpha} x_{t-1}) - \hat{\lambda} H(t)}{\hat{\alpha}(1-\hat{\alpha})} \right] = 0$$

$$\Rightarrow \sum_{t=1}^n x_t - \hat{\alpha} \sum_{t=1}^n x_{t-1} - \hat{\lambda} \sum_{t=1}^n H(t) = 0 \quad (2.16)$$

sustituyendo el resultado de S_λ tenemos:

$$\Rightarrow \sum_{t=1}^n x_t - \hat{\alpha} \sum_{t=1}^n x_{t-1} - \hat{\lambda} n = 0 \Rightarrow \hat{\lambda} = \frac{1}{n} \left[\sum_{t=1}^n x_t - \hat{\alpha} \sum_{t=1}^n x_{t-1} \right] \quad \blacksquare$$

Tomando una muestra de (n) observaciones, notamos que no encontramos los estimadores de máxima verosimilitud para α , de una forma explícita, por consiguiente no se encuentra el estimador de λ . Por lo que para encontrar los estimadores de α y λ usando la ecuación (2.16), se resuelve usando un método numérico.

Ahora, si tomamos $(n+1)$ observaciones para encontrar los estimadores de α y λ , usando la ecuación (2.9):

$$\begin{aligned} L(x, \alpha, \lambda) &= \left(\prod_{t=1}^n P_t(x_t) \right) \frac{[\lambda/(1-\alpha)]^{x_0} e^{-[\lambda/(1-\alpha)]}}{x_0!} \\ S_\lambda^0 &= \frac{\partial \ln L}{\partial \lambda} = S_\lambda + \frac{x_0}{\lambda} - \frac{1}{(1-\alpha)} \\ S_\alpha^0 &= \frac{\partial \ln L}{\partial \alpha} = S_\alpha + \frac{1}{1-\alpha} \left(x_0 - \frac{\lambda}{(1-\alpha)} \right) \end{aligned}$$

aplicando logaritmo natural a la ecuación (2.9), al igual que para S_λ :

$$\ln L(x, \alpha, \lambda) = \ln \left(\prod_{t=1}^n P_t(x_t) \right) + x_0 \ln \lambda - x_0 \ln(1-\alpha) - \ln x_0! - \frac{\lambda}{(1-\alpha)}$$

derivando esta expresión con respecto a λ , tenemos:

$$\frac{\partial \ln L(x, \alpha, \lambda)}{\partial \lambda} = \frac{\partial \ln \left(\prod_{t=1}^n P_t(x_t) \right)}{\partial \lambda} + \frac{x_0}{\lambda} - \frac{1}{(1-\alpha)} = S_\lambda + \frac{x_0}{\lambda} - \frac{1}{(1-\alpha)} \quad \blacksquare$$

derivando esta expresión con respecto a α , tenemos:

$$\begin{aligned} \frac{\partial \ln L(x, \alpha, \lambda)}{\partial \alpha} &= \frac{\partial \ln \left(\prod_{t=1}^n P_t(x_t) \right)}{\partial \alpha} + \frac{x_0}{(1-\alpha)} + \frac{\lambda}{(1-\alpha)^2} \\ &= S_\alpha + \frac{1}{(1-\alpha)} \left(x_0 + \frac{\lambda}{(1-\alpha)^2} \right) \quad \blacksquare \end{aligned}$$

observamos que las expresiones S_λ , S_α son las encontradas para los estimadores de λ y α usando (n) observaciones.

Igualando a cero S_λ^0 y S_α^0 , se tiene:

$$\begin{aligned} S_\lambda^0 &= S_\lambda + \frac{x_0}{\lambda} - \frac{1}{(1-\alpha)} = 0 \\ S_\alpha^0 &= S_\alpha + \frac{1}{(1-\alpha)} \left(x_0 + \frac{\lambda}{(1-\alpha)^2} \right) = 0 \end{aligned}$$

Al observar S_λ^0 y S_α^0 para una muestra de $(n + 1)$ observaciones, se verifica que no se pueden encontrar los estimadores de máxima verosimilitud para λ y α de manera explícita, ya que dependen de S_λ y S_α , ecuaciones que no se pueden resolver de manera explícita para (n) observaciones. Por lo tanto, para encontrar los estimadores de λ y α para $(n + 1)$ observaciones, se deben también utilizar métodos numéricos. ■

2.2. Estimación de Parámetros del Modelo INAR(2) con marginal de Poisson.

Recordando, el modelo INAR(2) está dado por:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \epsilon_t$$

La Función de Autocorrelación está dada por:

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2)$$

2.2.1. Estimadores de Yule-Walker.

Para un modelo autorregresivo de segundo orden, el conjunto de parámetros a estimar son $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda})$, auxiliándonos de la función de autocovarianza $\gamma(k) = \alpha_1 \gamma(k-1) + \alpha_2 \gamma(k-2)$, el método de momentos, que consiste en reemplazar la función de autocovarianza teórica por la función de autocovarianza muestral, por lo que de forma matricial tenemos:

$$\begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) \end{bmatrix} \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} = \begin{bmatrix} \hat{\gamma}(1) \\ \hat{\gamma}(2) \end{bmatrix}$$

La función de autocovarianza muestral $\hat{\gamma}(k)$ de x es:

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X}) \quad (2.17)$$

donde \bar{X} es la media muestral.

Por lo que los estimadores para $(\hat{\alpha}_1, \hat{\alpha}_2)$ se obtienen resolviendo el sistema de ecuaciones del producto matricial anterior:

$$\gamma(0)\hat{\alpha}_1 + \gamma(1)\hat{\alpha}_2 = \gamma(1) \quad (2.18)$$

$$\gamma(1)\hat{\alpha}_1 + \gamma(0)\hat{\alpha}_2 = \gamma(2) \quad (2.19)$$

Despejando $\hat{\alpha}_1$ de (2.18), se tiene:

$$\begin{aligned} \gamma(0)\hat{\alpha}_1 &= \gamma(1) - \gamma(1)\hat{\alpha}_2 \\ \hat{\alpha}_1 &= \frac{\gamma(1)(1 - \hat{\alpha}_2)}{\gamma(0)} = \rho(1)(1 - \hat{\alpha}_2) \end{aligned} \quad (2.20)$$

Despejando $\hat{\alpha}_2$ de (2.19), se tiene:

$$\begin{aligned}\gamma(0)\hat{\alpha}_2 &= \gamma(2) - \gamma(1)\hat{\alpha}_1 \\ \hat{\alpha}_2 &= \frac{\gamma(2) - \gamma(1)\hat{\alpha}_1}{\gamma(0)} = \rho(2) - \rho(1)\hat{\alpha}_1\end{aligned}\tag{2.21}$$

sustituyendo (2.20) en (2.21), se tiene:

$$\begin{aligned}\hat{\alpha}_2 &= \rho(2) - \rho(1)(\rho(1)(1 - \hat{\alpha}_2)) = \rho(2) - \rho^2(1) + \rho^2(1)\hat{\alpha}_2 \\ \hat{\alpha}_2(1 - \rho^2(1)) &= \rho(2) - \rho^2(1) \\ \hat{\alpha}_2 &= \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)} \blacksquare\end{aligned}$$

ahora, sustituyendo esta última expresión en (2.20), tenemos:

$$\hat{\alpha}_1 = \frac{\gamma(1)(1 - \hat{\alpha}_2)}{\gamma(0)} = \rho(1) \left(1 - \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)}\right) \blacksquare$$

Ahora el estimador para λ se puede obtener:

$$\hat{\lambda} = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha}_1 * X_{t-1} - \hat{\alpha}_2 * X_{t-2})\tag{2.22}$$

En efecto:

$$\begin{aligned}\epsilon_t &= X_t - \alpha_1 * X_{t-1} - \alpha_2 * X_{t-2} \\ E(\epsilon_t) &= E(X_t) - \alpha_1 E(X_{t-1}) - \alpha_2 E(X_{t-2}) \\ E(\widehat{\epsilon}_t) &= \frac{1}{n} \left(\sum_{t=1}^n X_t - \hat{\alpha}_1 \sum_{t=1}^n X_{t-1} - \hat{\alpha}_2 \sum_{t=1}^n X_{t-2} \right) \\ &= \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha}_1 X_{t-1} - \hat{\alpha}_2 X_{t-2})\end{aligned}$$

ya que ϵ_t tiene una distribución de Poisson de parámetro λ . \blacksquare

2.2.2. Estimadores Mínimos Cuadrados Condicionales.

Para encontrar los estimadores de α_1 , α_2 y λ de un modelo INAR(2), mediante el método de mínimos cuadrados condicionales, hacemos uso de una expresión (similar para el modelo INAR(1)), donde la media condicional de X_t , dado X_{t-1} y X_{t-2} es:

$$E[X_t/(X_{t-1}, X_{t-2})] = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda \quad (2.23)$$

La estimación inicia cuando aplicamos la teoría desarrollada por Klimko y Nelson (1978) se basa en la minimización de la suma de desviaciones cuadradas acerca de la esperanza condicional. Así los estimadores de mínimos cuadrados condicionales (CLS) para α_1 , α_2 y λ son los valores que minimizan:

$$Q_n(\alpha_1, \alpha_2, \lambda) = \sum_{t=2}^n (X_t - E[X_t/(X_{t-1}, X_{t-2})])^2 \quad (2.24)$$

con respecto a $(\alpha_1, \alpha_2, \lambda)$.

Derivando la ecuación (2.24) con respecto a α_1 , α_2 y λ , se tiene:

$$\hat{\alpha}_1 = \frac{\left[\frac{\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]}{\left[\frac{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] - \left[\frac{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]} \quad (2.25)$$

$$\hat{\alpha}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] - \left[\frac{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right]} \quad (2.26)$$

y

$$\hat{\lambda} = \bar{X}_t - \hat{\alpha}_1 \bar{X}_{t-1} - \hat{\alpha}_2 \bar{X}_{t-2} \quad (2.27)$$

Sustituyendo $\hat{\alpha}_1$ y $\hat{\alpha}_2$, respectivamente.

Verificación:

De la ecuación (2.24) y desarrollando el binomio, se tiene:

$$\begin{aligned} Q_n(\alpha_1, \alpha_2, \lambda) &= \sum_{t=2}^n (X_t - E[X_t/(X_{t-1}, X_{t-2})])^2 \\ &= \sum_{t=2}^n (X_t - (\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda))^2 \\ &= \sum_{t=2}^n (X_t^2 - 2X_t(\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda) + (\alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \lambda)^2) \\ &= \sum_{t=2}^n (X_t^2 - 2\alpha_1 X_t X_{t-1} - 2\alpha_2 X_t X_{t-2} - 2\lambda X_t \\ &\quad + \alpha_1^2 X_{t-1}^2 + 2\alpha_1 \alpha_2 X_{t-1} X_{t-2} + \alpha_2^2 X_{t-2}^2 + 2\alpha_1 \lambda X_{t-1} + 2\alpha_2 \lambda X_{t-2} + \lambda^2) \end{aligned}$$

derivando $Q_n(\alpha_1, \alpha_2, \lambda)$ con respecto a α_1 , α_2 y λ , tenemos:

$$\frac{\partial Q_n(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_1} = \sum_{t=2}^n (-2X_t X_{t-1} + 2\alpha_1 X_{t-1}^2 + 2\alpha_2 X_{t-1} X_{t-2} + 2\lambda X_{t-1}) = 0 \quad (2.28)$$

$$\frac{\partial Q_n(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_2} = \sum_{t=2}^n (-2X_t X_{t-2} + 2\alpha_1 X_{t-1} X_{t-2} + 2\alpha_2 X_{t-2}^2 + 2\lambda X_{t-1}) = 0 \quad (2.29)$$

$$\frac{\partial Q_n(\alpha_1, \alpha_2, \lambda)}{\partial \lambda} = \sum_{t=2}^n (-2X_t + 2\alpha_1 X_{t-1} + 2\alpha_2 X_{t-2} + 2\lambda) = 0 \quad (2.30)$$

de la derivada de $Q_n(\alpha_1, \alpha_2, \lambda)$ con respecto a λ se tiene:

$$\hat{\lambda} = \frac{1}{n-1} \left[\sum_{t=2}^n X_t - \hat{\alpha}_1 \sum_{t=2}^n X_{t-1} - \hat{\alpha}_2 \sum_{t=2}^n X_{t-2} \right]$$

$$\hat{\lambda} = \bar{X}_t - \hat{\alpha}_1 \bar{X}_{t-1} - \hat{\alpha}_2 \bar{X}_{t-2} \quad \blacksquare$$

sustituyendo el estimador de $\hat{\lambda}$ en la ecuación (2.28) tenemos para $\hat{\alpha}_1$:

$$\begin{aligned}
\frac{\partial Q_n(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_1} &= \sum_{t=2}^n (-2X_t X_{t-1} + 2\alpha_1 X_{t-1}^2 + 2\alpha_2 X_{t-1} X_{t-2} + 2\lambda X_{t-1}) = 0 \\
\Rightarrow -\sum_{t=2}^n X_t X_{t-1} + \hat{\alpha}_1 \sum_{t=2}^n X_{t-1}^2 + \hat{\alpha}_2 \sum_{t=2}^n X_{t-1} X_{t-2} + \hat{\lambda} \sum_{t=2}^n X_{t-1} &= 0 \\
\Rightarrow -\sum_{t=2}^n X_t X_{t-1} + \hat{\alpha}_1 \sum_{t=2}^n X_{t-1}^2 + \hat{\alpha}_2 \sum_{t=2}^n X_{t-1} X_{t-2} \\
&\quad + (\bar{X}_t - \hat{\alpha}_1 \bar{X}_{t-1} - \hat{\alpha}_2 \bar{X}_{t-2}) \sum_{t=2}^n X_{t-1} = 0
\end{aligned}$$

despejando $\hat{\alpha}_1$ de éste último resultado:

$$\begin{aligned}
\hat{\alpha}_1 &= \frac{\left[\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1} \right] + \hat{\alpha}_2 \left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]} \\
&= \frac{\left[\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]} + \hat{\alpha}_2 \frac{\left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]} \quad (2.31)
\end{aligned}$$

ahora, sustituyendo el estimador de λ en la ecuación (2.29) tenemos para $\hat{\alpha}_2$:

$$\begin{aligned}
\frac{\partial Q_n(\alpha_1, \alpha_2, \lambda)}{\partial \alpha_2} &= \sum_{t=2}^n (-2X_t X_{t-2} + 2\alpha_1 X_{t-1} X_{t-2} + 2\alpha_2 X_{t-2}^2 + 2\lambda X_{t-1}) = 0 \\
\Rightarrow -\sum_{t=2}^n X_t X_{t-2} + \hat{\alpha}_1 \sum_{t=2}^n X_{t-1} X_{t-2} + \hat{\alpha}_2 \sum_{t=2}^n X_{t-2}^2 + \hat{\lambda} \sum_{t=2}^n X_{t-1} &= 0 \\
\Rightarrow -\sum_{t=2}^n X_t X_{t-2} + \hat{\alpha}_2 \sum_{t=2}^n X_{t-2}^2 + \hat{\alpha}_1 \sum_{t=2}^n X_{t-1} X_{t-2} \\
&\quad + (\bar{X}_t - \hat{\alpha}_1 \bar{X}_{t-1} - \hat{\alpha}_2 \bar{X}_{t-2}) \sum_{t=2}^n X_{t-1} = 0
\end{aligned}$$

despejando $\hat{\alpha}_2$ de este último resultado:

$$\begin{aligned}\hat{\alpha}_2 &= \frac{\left[\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2} \right] + \hat{\alpha}_1 \left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} \\ &= \frac{\left[\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} + \hat{\alpha}_1 \frac{\left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} \quad (2.32)\end{aligned}$$

hasta el momento, hemos logrado que las expresiones de $\hat{\alpha}_1$ y $\hat{\alpha}_2$, no dependan de $\hat{\lambda}$, pero dependen de $\hat{\alpha}_2$ y $\hat{\alpha}_1$ respectivamente, por lo que sustituyendo la ecuación (2.32) en la ecuación (2.31), tenemos:

$$\begin{aligned}\hat{\alpha}_1 &= \frac{\left[\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]} + \frac{\left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]} \\ &\times \left(\frac{\left[\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} + \hat{\alpha}_1 \frac{\left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} \right)\end{aligned}$$

sabemos que:

$$\begin{aligned}\left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right] &= \left[\frac{1}{n-1} \sum_{t=2}^n X_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right] \\ &= \left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right] \quad (2.33)\end{aligned}$$

usando este resultado, para encontrar el estimador de α_1 se tiene:

$$\hat{\alpha}_1 \left(\frac{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} - \frac{\left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]^2}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]} \right) =$$

$$\left[\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1} \right] + \left[\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2} \right] \frac{\left[\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2} \right]}{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right]}$$

por lo que finalmente, la expresión de $\hat{\alpha}_1$ es:

$$\hat{\alpha}_1 = \frac{\left[\frac{\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]}{\left[\frac{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] - \left[\frac{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]} \quad \blacksquare$$

Ahora, para encontrar el estimador de $\hat{\alpha}_2$, se sigue un proceso similar para $\hat{\alpha}_1$, por lo que sustituyendo la ecuación (2.31) en la ecuación (2.32), tenemos:

$$\hat{\alpha}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right] + \left[\frac{\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]}{\left(\frac{\left[\frac{\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right] + \hat{\alpha}_2 \left[\frac{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}}{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}} \right]} \right)}$$

usando el resultado de la ecuación (2.33), tenemos:

$$\hat{\alpha}_2 \left(\frac{\left[\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2} \right] - \frac{\left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right]^2}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]}}{\left[\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2} \right] + \left[\bar{X}_{t-1} \sum_{t=2}^n X_{t-2} - \sum_{t=2}^n X_{t-1} X_{t-2} \right] \frac{\left[\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1} \right]}{\left[\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1} \right]}} \right) =$$

Por lo que la expresión de $\hat{\alpha}_2$ es:

$$\hat{\alpha}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t X_{t-2} - \bar{X}_t \sum_{t=2}^n X_{t-2}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t X_{t-1} - \bar{X}_t \sum_{t=2}^n X_{t-1}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n X_{t-2}^2 - \bar{X}_{t-2} \sum_{t=2}^n X_{t-2}}{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}} \right] - \left[\frac{\bar{X}_{t-2} \sum_{t=2}^n X_{t-1} - \sum_{t=2}^n X_{t-1} X_{t-2}}{\sum_{t=2}^n X_{t-1}^2 - \bar{X}_{t-1} \sum_{t=2}^n X_{t-1}} \right]} \quad \blacksquare$$

Finalmente, sustituyendo las expresiones encontradas de $\hat{\alpha}_1$ y $\hat{\alpha}_2$ en la ecuación (2.27), se obtiene el estimador de $\hat{\lambda}$. \blacksquare

Se observa que usando el método de mínimos cuadrados condicionales, las expresiones encontradas para $\hat{\alpha}_1$ y $\hat{\alpha}_2$ son muy complejas, por lo que también se vuelve compleja la expresión para $\hat{\lambda}$ y esa complejidad aumenta con el orden del modelo.

2.2.3. Estimadores de Máxima Verosimilitud.

De igual forma que para encontrar los estimadores de un modelo INAR(1), la función de probabilidad (verosimilitud) de una muestra de $(n + 1)$ observaciones $\{X_0, X_1, \dots, X_t\}$ para el proceso INAR(2) puede escribirse como:

$$L(x, \alpha_1, \alpha_2, \lambda) = \left(\prod_{t=2}^n P_t(x_t) \right) \frac{[\lambda/(1 - \alpha_1 - \alpha_2)]^{x_0+x_1}}{x_0!x_1!} e^{-2\lambda/(1-\alpha_1-\alpha_2)} \quad (2.34)$$

En efecto:

La función de verosimilitud podemos escribirla como el producto de las funciones de probabilidad de los x_i para $i = 0, 1, \dots, t$.

$$L(x, \alpha, \lambda) = P_t(x_t) \cdot P_{t-1}(x_{t-1}) \dots P_1(x_1) \cdot P_0(x_0)$$

De la definición del modelo INAR(2) se tiene:

$$\begin{aligned} x_t &= \alpha_1 * x_{t-1} + \alpha_2 * x_{t-2} + \epsilon_t \\ x_0 &= \alpha_1 * x_{-1} + \alpha_2 * x_{-2} + \epsilon_0 = \epsilon_0 \\ x_1 &= \alpha_1 * x_0 + \alpha_2 * x_{-1} + \epsilon_1 = \alpha_1 * x_0 + \epsilon_1 = \alpha_1 * \epsilon_0 + \epsilon_1 \end{aligned}$$

dado que conocemos únicamente los sucesos desde 0 hasta t , además vemos que, ϵ_0 tendrá la misma distribución de x_0 ; la distribución de x_1 dependerá de la suma de las distribuciones de $\alpha_1 * \epsilon_0$ y ϵ_1 .

Al inicio de este apartado, se asume que ϵ_t tiene distribución de Poisson y sabemos que el valor esperado del proceso inicial estacionario es $E(x_0) = E(x_1) = \mu_\epsilon/(1 - \alpha_1 - \alpha_2)$:

$$\begin{aligned} P(x_0) &= \frac{[\lambda/(1 - \alpha_1 - \alpha_2)]^{x_0}}{x_0!} e^{-\lambda/(1-\alpha_1-\alpha_2)} \\ P(x_1) &= \frac{[\lambda/(1 - \alpha_1 - \alpha_2)]^{x_1}}{x_1!} e^{-\lambda/(1-\alpha_1-\alpha_2)} \end{aligned}$$

estos resultados, los sustituimos en $L(x, \alpha_1, \alpha_2, \lambda)$ y el producto de las distribuciones se tomará para $t = 2, 3, \dots, n$, de manera que:

$$L(x, \alpha_1, \alpha_2, \lambda) = \left(\prod_{t=2}^n P_t(x_t) \right) \frac{[\lambda/(1 - \alpha_1 - \alpha_2)]^{x_0+x_1}}{x_0!x_1!} e^{-2\lambda/(1-\alpha_1-\alpha_2)} \quad \blacksquare$$

Se puede verificar que:

$$P_t(X_t = x_t) = e^{-\lambda} \sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t-i-j}}{(x_t-i-j)!} \binom{x_{t-1}}{i} \alpha_1^i (1-\alpha_1)^{x_{t-1}-i} \binom{x_{t-2}}{j} \alpha_2^j (1-\alpha_2)^{x_{t-2}-j} \quad (2.35)$$

para $t = 2, 3, \dots, n$, $M = \min(x_{t-1}, x_t)$ y $N = \min(x_{t-2}, x_t - i)$.

Verificación:

Para verificar este resultado, utilizaremos la distribución de masa de tres variables.

Sea x_1, x_2 y x_3 variables aleatorias, entonces $x_1 + x_2 + x_3$ es también variable aleatoria. Si $x_1 + x_2 + x_3$ es una variable aleatoria discreta, su función de masa de probabilidad viene dada por:

$$P_{x_1+x_2+x_3}(y) = \sum_K \sum_{K'} P_{x_1}(i) P_{x_2}(j) P_{x_3}(y-i-j)$$

donde $y = x_1 + x_2 + x_3$ y $\forall i, j$ se cumple que $y - i - j \geq 0$ siendo i la i -ésima observación del conjunto K y j la j -ésima observación del conjunto K' .

Esta función de masa de probabilidad, está definida para todo K y K' donde se cumplen las condiciones:

- $P_{x_1}(i) > 0$
- $P_{x_2}(j) > 0$
- $P_{x_3}(y - i - j) > 0$

Sabemos que, $P_t(x_t) = P_t(x_t = \alpha_1 * x_{t-1} + \alpha_2 * x_{t-2} + \epsilon_t)$ es la distribución de la suma de tres v.a. discretas.

Por lo que tenemos:

$$P_{\alpha_1 * x_{t-1} + \alpha_2 * x_{t-2} + \epsilon_t}(x_t) = \sum_{i=0}^M \sum_{j=0}^N P_{\alpha_1 * x_{t-1}}(i) P_{\alpha_2 * x_{t-2}}(j) P_{\epsilon_t}(x_t - i - j)$$

se asume que ϵ_t tiene una distribución de Poisson de parámetro $\lambda > 0$, que está definido para todo $i \geq 0, j \geq 0$:

$$P_{\epsilon_t}(x_t - i - j) = \frac{\lambda^{x_t-i-j}}{(x_t-i-j)!} e^{-\lambda} \quad (2.36)$$

sabe que $(\alpha_1 * x_{t-1}/x_{t-1})$ tiene distribución Binomial con parámetros α_1 y x_{t-1} , entonces de $\epsilon_t = x_t - i - j$, se tiene que $\alpha * x_{t-1} = i$, por lo tanto:

$$P_{\alpha_1 * x_{t-1}}(i) = \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1}-i} \quad (2.37)$$

además se sabe que $(\alpha_2 * x_{t-2}/x_{t-2})$ tiene distribución Binomial con parámetros α_2 y x_{t-2} , entonces de $\epsilon_t = x_t - i - j$ y $\alpha * x_{t-1} = i$ se tiene que $\alpha_2 * x_{t-2} = j$, por lo tanto:

$$P_{\alpha_2 * x_{t-2}}(j) = \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2}-j} \quad (2.38)$$

sustituyendo las ecuaciones (2.36), (2.37) y (2.38) en (2.35) tenemos:

$$P_t(X_t = x_t) = e^{-\lambda} \sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x-i-j}}{(x-i-j)!} \binom{x_{t-1}}{i} \alpha^i (1 - \alpha)^{x_{t-1}-i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2}-j}$$

Esta última expresión es la ecuación (2.35) definida anteriormente, pero notamos que la sumatoria de esta expresión está definida desde $i = 0$ hasta $M = \min(x_{t-1}, x_t)$ (ver página 54) y $j = 0$ hasta $N = \min(x_{t-2}, x_t - i)$, para comprender $N = \min(x_{t-2}, x_t - i)$, hacemos uso de dos posibles resultados:

- Supongamos que $M = \min(x_{t-1}, x_t) = x_{t-1}$.

Entonces, la segunda sumatoria se define desde $j = 0$ hasta $N = \min(x_{t-2}, x_t - x_{t-1})$, de aquí obtenemos dos resultados:

- Si $N = \min(x_{t-2}, x_t - x_{t-1}) = x_{t-2}$, el orden en los datos de la sumatoria es $j = 0, 1, 2, \dots$. Notamos que el combinatorio $\binom{x_{t-2}}{j}$ y $(1 - \alpha_2)^{x_{t-2}-j}$ en la ecuación (2.35) está definido hasta x_{t-2} .

Por lo que si $N = \min(x_{t-2}, x_t - x_{t-1}) = x_{t-2}$, entonces la sumatoria está definida para $j = 0, 1, 2, \dots, x_{t-2}$.

- Si $N = \min(x_{t-2}, x_t - x_{t-1}) = x_t - x_{t-1}$, el orden en los datos de la sumatoria es $j = 0, 1, 2, \dots$. Notamos que el factorial $(x_t - i - j)! = (x_t - x_{t-1} - (x_t - x_{t-1}))!$ y $\lambda^{x_i - x_{t-1} - (x_t - x_{t-1})}$ en la ecuación (2.35) está definido hasta $x_t - x_{t-1}$.

Por lo que si $N = \min(x_{t-2}, x_t - x_{t-1}) = x_t - x_{t-1}$, entonces la sumatoria está definida para $j = 0, 1, 2, \dots, x_t - x_{t-1}$

- Supongamos que $M = \min(x_{t-1}, x_t) = x_t$.
- Si $N = \min(x_{t-2}, x_t - x_t) = 0$, notamos que en este caso, la segunda sumatoria no se aplica.

De estos resultados se comprueba que, la segunda sumatoria está definida desde $j = 0$ hasta $N = \min(x_{t-2}, x_t - i)$ si la primer sumatoria está definida para $i = 0$ hasta $M = \min(x_{t-1}, x_t)$. ■

Para α_1 , α_2 y λ dado x_0 y x_1 , consideraremos los primeros estimadores condicionales de máxima verosimilitud, en este caso la función condicional de verosimilitud se define:

$$L(x, \alpha_1, \alpha_2, \lambda/x_0, x_1) = \prod_{t=2}^n P_{x_t/x_{t-2}, x_{t-1}}(x_t) = \prod_{t=2}^n P_t(x_t) \tag{2.39}$$

Usando los resultados obtenidos por Sportt (1983) de las derivadas de los logaritmos de las estimaciones de la función condicional de verosimilitud para α_1 , α_2 y λ , tomando $(n - 1)$ observaciones, tenemos:

$$S_\lambda = \sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \lambda} = \frac{1}{\lambda} \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) \tag{2.40}$$

$$S_{\alpha_1} = \sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha_1} = \frac{1}{\alpha_1(1 - \alpha_1)} \sum_{t=2}^n \left(\frac{P_t(x_t)i}{P_t(x_t)} - \alpha_1 x_{t-1} \right) \tag{2.41}$$

$$S_{\alpha_2} = \sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha_2} = \frac{1}{\alpha_2(1 - \alpha_2)} \sum_{t=2}^n \left(\frac{P_t(x_t)j}{P_t(x_t)} - \alpha_2 x_{t-2} \right) \tag{2.42}$$

En efecto:

Para verificar el resultado de S_λ , tenemos:

De la ecuación (2.35) se tiene:

$$P_t(x_t) = e^{-\lambda} \sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t-i-j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1}-i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2}-j}$$

Aplicando logaritmo natural a $P_t(x_t)$ se tiene:

$$\ln P_t(x_t) = -\lambda + \ln \left[\sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t-i-j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1}-i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2}-j} \right]$$

derivando esta última expresión con respecto a λ :

$$\frac{\partial \ln P_t(x_t)}{\partial \lambda} = -1 + \frac{\left[\sum_{i=0}^M \sum_{j=0}^N \frac{(x_t - i - j) \lambda^{x_t - i - j - 1}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1} - i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]}{\left[\sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t - i - j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1} - i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]}$$

si la expresión anterior se multiplica y divide por $\lambda e^{-\lambda}$, se obtiene $P_t(x_t)$ en el denominador:

$$\begin{aligned} \frac{\partial \ln P_t(x_t)}{\partial \lambda} &= -1 \\ &+ \frac{e^{-\lambda}}{\lambda} \frac{\left[\sum_{i=0}^M \sum_{j=0}^N \frac{(x_t - i - j) \lambda \lambda^{x_t - i - j - 1}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1} - i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]}{P_t(x_t)} \\ &= -1 + \frac{1}{\lambda} \frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} \end{aligned}$$

de esta expresión resulta fácil ver que:

$$\sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \lambda} = \frac{1}{\lambda} \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) \quad \blacksquare$$

Ahora, para encontrar la expresión S_{α_1} , partimos de $\ln P_t(x_t)$, expresión usada para encontrar S_λ :

$$\ln P_t(x_t) = -\lambda + \ln \left[\sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t - i - j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1} - i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]$$

derivando esta expresión con respecto a α_1 :

$$\begin{aligned} \frac{\partial \ln P_t(x_t)}{\partial \alpha_1} &= \frac{\left[\sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t - i - j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]}{\left[\sum_{i=0}^M \sum_{j=0}^N \frac{\lambda^{x_t - i - j}}{(x_t - i - j)!} \binom{x_{t-1}}{i} \alpha_1^i (1 - \alpha_1)^{x_{t-1} - i} \binom{x_{t-2}}{j} \alpha_2^j (1 - \alpha_2)^{x_{t-2} - j} \right]} \\ &\times \left(i \alpha_1^{i-1} (1 - \alpha_1)^{x_{t-1} - i} - \alpha_1^i (x_{t-1} - i) (1 - \alpha_1)^{x_{t-1} - i - 1} \right) \end{aligned}$$

si multiplicamos y dividimos esta expresión por $e^{-\lambda}$ y obtenemos $(\alpha_1^i(1 - \alpha_1)^{x_{t-1}-i})$ de factor común, observamos que se obtiene $P_t(x_t)$ en el numerador y denominador:

$$\begin{aligned} \frac{\partial \ln P_t(x_t)}{\partial \alpha_1} &= \frac{P_t(x_t) (i\alpha_1^{-1} - (x_{t-1} - i)(1 - \alpha_1)^{-1})}{P_t(x_t)} \\ &= \frac{P_t(x_t) \left(\frac{i}{\alpha_1} - \frac{(x_{t-1} - i)}{(1 - \alpha_1)} \right)}{P_t(x_t)} = \frac{P_t(x_t) \left(\frac{i(1 - \alpha_1) - \alpha_1(x_{t-1} - i)}{\alpha_1(1 - \alpha_1)} \right)}{P_t(x_t)} \\ &= \frac{1}{\alpha_1(1 - \alpha_1)} \frac{P_t(x_t)(i - \alpha_1 x_{t-1})}{P_t(x_t)} \end{aligned}$$

aplicando la sumatoria a la expresión anterior, se tiene:

$$\sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha_1} = \frac{1}{\alpha_1(1 - \alpha_1)} \sum_{t=2}^n \left(\frac{P_t(x_t)i}{P_t(x_t)} - \alpha_1 x_{t-1} \right) \quad \blacksquare$$

Para encontrar la expresión S_{α_2} , se siguen los mismos pasos utilizados para encontrar S_{α_1} , en este caso, se deriva $\ln P_t(x_t)$ con respecto a α_2 , por lo que se obtiene:

$$\sum_{t=2}^n \frac{\partial \ln P_t(x_t)}{\partial \alpha_2} = \frac{1}{\alpha_2(1 - \alpha_2)} \sum_{t=2}^n \left(\frac{P_t(x_t)j}{P_t(x_t)} - \alpha_2 x_{t-2} \right) \quad \blacksquare$$

Realizando algunas operaciones algebraicas en S_{α_1} y S_{α_2} se tiene:

$$\sum_{t=2}^n \left(\frac{P_t(x_t)(i + j)}{P_t(x_t)} \right) = \alpha_1 \sum_{t=2}^n x_{t-1} + \alpha_2 \sum_{t=2}^n x_{t-2} \quad (2.43)$$

En efecto:

Haciendo $S_\lambda = 0$, $S_{\alpha_1} = 0$ y $S_{\alpha_2} = 0$, respectivamente tenemos:

$$\begin{aligned} S_\lambda &= \frac{1}{\lambda} \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) = \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) = 0 \\ &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} = \sum_{t=2}^n \lambda \\ S_{\alpha_1} &= \frac{1}{\alpha_1(1 - \alpha_1)} \sum_{t=2}^n \left(\frac{P_t(x_t)i}{P_t(x_t)} - \alpha_1 x_{t-1} \right) = \sum_{t=2}^n \left(\frac{P_t(x_t)i}{P_t(x_t)} - \alpha_1 x_{t-1} \right) = 0 \\ &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)i}{P_t(x_t)} = \sum_{t=2}^n \alpha_1 x_{t-1} \\ S_{\alpha_2} &= \frac{1}{\alpha_2(1 - \alpha_2)} \sum_{t=2}^n \left(\frac{P_t(x_t)j}{P_t(x_t)} - \alpha_2 x_{t-2} \right) = \sum_{t=2}^n \left(\frac{P_t(x_t)j}{P_t(x_t)} - \alpha_2 x_{t-2} \right) = 0 \\ &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)j}{P_t(x_t)} = \sum_{t=2}^n \alpha_2 x_{t-2} \end{aligned}$$

como $S_{\alpha_1} = 0$ y $-S_{\alpha_2} = 0$, entonces:

$$\begin{aligned} S_{\alpha_1} = -S_{\alpha_2} &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)i}{P_t(x_t)} - \alpha_1 \sum_{t=2}^n x_{t-1} = -\sum_{t=2}^n \frac{P_t(x_t)j}{P_t(x_t)} + \alpha_2 \sum_{t=2}^n x_{t-2} \\ &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)(i+j)}{P_t(x_t)} = \alpha_1 \sum_{t=2}^n x_{t-1} + \alpha_2 \sum_{t=2}^n x_{t-2} \end{aligned}$$

Ahora, si usamos la expresión $S_\lambda = 0$ y la ecuación (2.43), podemos encontrar un estimador de λ :

$$\begin{aligned} S_\lambda &= \frac{1}{\lambda} \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) = \sum_{t=2}^n \left(\frac{P_t(x_t)(x_t - i - j)}{P_t(x_t)} - \lambda \right) = 0 \\ &\Rightarrow \sum_{t=2}^n \frac{P_t(x_t)(x_t)}{P_t(x_t)} - \sum_{t=2}^n \frac{P_t(x_t)(i+j)}{P_t(x_t)} - \sum_{t=2}^n \lambda = 0 \end{aligned}$$

sustituyendo la expresión (2.43) en este último resultado:

$$\sum_{t=2}^n \frac{P_t(x_t)(x_t)}{P_t(x_t)} - \alpha_1 \sum_{t=2}^n x_{t-1} - \alpha_2 \sum_{t=2}^n x_{t-2} - (n-1)\lambda = 0$$

despejando λ , tenemos:

$$\begin{aligned} (n-1)\lambda &= \sum_{t=2}^n x_t - \alpha_1 \sum_{t=2}^n x_{t-1} - \alpha_2 \sum_{t=2}^n x_{t-2} \\ \lambda &= \frac{1}{(n-1)} \left[\sum_{t=2}^n x_t - \alpha_1 \sum_{t=2}^n x_{t-1} - \alpha_2 \sum_{t=2}^n x_{t-2} \right] \end{aligned}$$

Observemos que se puede obtener un estimador de forma explícita para λ , pero no se obtienen estimadores para α_1 y α_2 de manera que no dependa de otro parámetro, esto se observa de $S_\lambda = 0$, $S_{\alpha_1} = 0$ y $S_{\alpha_2} = 0$ que se expresan en término de $P_t(x_t)$ y este, depende de α_1 y α_2 .

Por lo que para encontrar los estimadores de los parámetros usando el método de máxima verosimilitud condicional, para $(n-1)$ observaciones se utiliza un método numérico. ■

2.3. Estimación de Parámetros del Modelo INAR(p) con marginal de Poisson.

Recordando, que el modelo INAR(p) está dado por:

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \dots + \alpha_p * X_{t-p} + \epsilon_t$$

La Función de Autocorrelación está dada por:

$$\rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \dots + \alpha_p \rho(k-p)$$

2.3.1. Estimadores de Yule-Walker.

Usando los resultados propuestos por Al-Osh y Alzaid (1987), Du y Li (1991) y Marques da Silva (2005), en el que generalizan el método de Yule-Walker (método de momentos), el que consiste en reemplazar la función de autocovarianza teórica por la función de autocovarianza muestral, es decir:

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

Por lo que los estimadores de Yule-Walker para $\alpha_1, \dots, \alpha_p$ se obtienen resolviendo el sistema de ecuaciones lineales en el orden de los parámetros del sistema anterior.

Donde:

$$\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})$$

y

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t$$

Ahora el estimador para λ se puede obtener:

$$\hat{\lambda} = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\alpha}_1 * X_{t-1} - \hat{\alpha}_2 * X_{t-2} - \dots - \hat{\alpha}_p * X_{t-p}) \tag{2.44}$$

En efecto:

$$\begin{aligned}\epsilon_t &= X_t - \alpha_1 * X_{t-1} - \alpha_2 * X_{t-2} - \dots - \alpha_p * X_{t-p} \\ E(\epsilon_t) &= E(X_t) - \alpha_1 E(X_{t-1}) - \alpha_2 E(X_{t-2}) - \dots - \alpha_p E(X_{t-p})\end{aligned}$$

ahora, un estimador razonable para λ es:

$$\begin{aligned}\widehat{E(\epsilon_t)} &= \frac{1}{n} \left(\sum_{t=1}^n X_t - \widehat{\alpha}_1 \sum_{t=1}^n X_{t-1} - \widehat{\alpha}_2 \sum_{t=1}^n X_{t-2} - \dots - \widehat{\alpha}_p \sum_{t=1}^n X_{t-p} \right) \\ &= \frac{1}{n} \sum_{t=1}^n (X_t - \widehat{\alpha}_1 X_{t-1} - \widehat{\alpha}_2 X_{t-2} - \dots - \widehat{\alpha}_p X_{t-p}) \quad \blacksquare\end{aligned}$$

ya que ϵ_t tiene una distribución de Poisson de parámetro λ .

2.3.2. Estimadores Mínimos Cuadrados Condicionales.

Se presenta un resumen de la generalización de dicho método que fue estudiado por Du y Li (1991) y Funes (2001).

En el modelo INAR(p) se tiene que dado $(X_{t-1}, X_{t-2}, \dots, X_{t-p})$ y ϵ_t , X_t sigue siendo v.a. entonces:

$$\begin{aligned} E(X_t / (X_{t-1}, X_{t-2}, \dots, X_{t-p})) &= \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \lambda \\ &= g(\theta, (X_{t-1}, X_{t-2}, \dots, X_{t-p})) \end{aligned}$$

donde $\theta = (\alpha_1, \alpha_2, \dots, \alpha_p, \lambda)^T$ es el conjunto de parámetros a estimar. 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:

$$\sum_{t=p}^n (X_t - E(X_t / (X_{t-1}, X_{t-2}, \dots, X_{t-p})))^2 \tag{2.45}$$

Diferenciando la ecuación (2.45) con respecto a θ_i con $i = 1, 2, \dots, (p+1)$ e igualando a cero se obtienen los estimadores de mínimos cuadrados condicionales de $\hat{\theta}_i$. ■

Concluyendo con la etapa de la estimación de parámetros de los modelos INAR, se vuelve complicado presentar una generalización del método de máxima verosimilitud de un modelo INAR(p), por lo que observando la estimación de un modelo INAR(1) e INAR(2), se puede seguir una idea de este proceso. Para ello, hay que tener en cuenta que la probabilidad condicional de X_t dado $X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_{t-p} = x_{t-p}$ es la convolución de la distribución de ϵ_t y las binomiales con parámetros $(\alpha_1, x_{t-1}), (\alpha_2, x_{t-2}), \dots, (\alpha_p, x_{t-p})$ respectivamente.

2.4. Estimación de Parámetros del Modelo INMA(1) con marginal de Poisson.

Recordemos que el modelo INMA(1) está dado por:

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

2.4.1. Estimadores Mínimos Cuadrados Condicionales.

Al igual que para un modelo INAR(1), para obtener los estimadores de β y λ de un modelo INMA(1), mediante los estimadores de mínimos cuadrados condicionales, hacemos uso de la expresión, donde la media condicional de X_t , dado ϵ_{t-1} es:

$$E[X_t/\epsilon_{t-1}] = \beta\epsilon_{t-1} + \lambda \quad (2.46)$$

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

$$Q_n(\beta, \lambda) = \sum_{t=1}^n (X_t - E[X_t/\epsilon_{t-1}])^2 \quad (2.47)$$

con respecto a β y λ .

Derivando (2.47) con respecto a β y λ se tiene:

$$\hat{\beta} = \frac{n \sum_{t=1}^n X_t \epsilon_{t-1} - \sum_{t=1}^n X_t \sum_{t=1}^n \epsilon_{t-1}}{n \sum_{t=1}^n \epsilon_{t-1}^2 - \left(\sum_{t=1}^n \epsilon_{t-1} \right)^2} \quad (2.48)$$

y

$$\hat{\lambda} = \frac{1}{n} \sum_{t=1}^n (X_t - \hat{\beta} \epsilon_{t-1}) \quad (2.49)$$

Verificación:

De la ecuación (2.47), se tiene:

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

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

$$\frac{\partial Q_n(\beta, \lambda)}{\partial \beta} = \sum_{t=1}^n (-2X_t\epsilon_{t-1} + 2\beta\epsilon_{t-1}^2 + 2\lambda\epsilon_{t-1}) = 0 \tag{2.50}$$

$$\frac{\partial Q_n(\beta, \lambda)}{\partial \lambda} = \sum_{t=1}^n (-2X_t + 2\beta\epsilon_{t-1} + 2\lambda) = 0 \tag{2.51}$$

de la derivada de $Q_n(\beta, \lambda)$ con respecto a λ se tiene que:

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

sustituyendo el estimador $\hat{\lambda}$ en la ecuación (2.50) tenemos para $\hat{\beta}$:

$$\begin{aligned}
 \frac{\partial Q_n(\beta, \lambda)}{\partial \beta} &= -2 \sum_{t=1}^n X_t\epsilon_{t-1} + 2\beta \sum_{t=1}^n \epsilon_{t-1}^2 + 2\lambda \sum_{t=1}^n \epsilon_{t-1} = 0 \\
 \Rightarrow & - \sum_{t=1}^n X_t\epsilon_{t-1} + \hat{\beta} \sum_{t=1}^n \epsilon_{t-1}^2 + \frac{1}{n} \left[\sum_{t=1}^n X_t - \hat{\beta} \sum_{t=1}^n \epsilon_{t-1} \right] \sum_{t=1}^n \epsilon_{t-1} = 0 \\
 \hat{\beta} &= \frac{n \sum_{t=1}^n X_t\epsilon_{t-1} - \sum_{t=1}^n X_t \sum_{t=1}^n \epsilon_{t-1}}{n \sum_{t=1}^n \epsilon_{t-1}^2 - \left(\sum_{t=1}^n \epsilon_{t-1} \right)^2} \quad \blacksquare
 \end{aligned}$$

2.5. Estimación de Parámetros del Modelo INMA(2) con marginal de Poisson.

Recordemos que el modelo INMA(2) está dado por:

$$X_t = \beta_1 * \epsilon_{t-1} + \beta_2 * \epsilon_{t-2} + \epsilon_t$$

2.5.1. Estimadores Mínimos Cuadrados Condicionales.

Siguiendo un proceso similar para los modelo INAR(2), los estimadores de β_1 , β_2 y λ de un modelo INMA(2), mediante los estimadores de mínimos cuadrados condicionales, se obtienen usando la media condicional de X_t , dado ϵ_{t-1} y ϵ_{t-2} :

$$E[X_t/(\epsilon_{t-1}, \epsilon_{t-2})] = \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \lambda \quad (2.52)$$

Como ya sabemos, la estimación se basa en la minimización de la suma de desviaciones cuadradas acerca de la esperanza condicional. Por lo que los estimadores de mínimos cuadrados condicionales (CLS) estimados para β_1 , β_2 y λ son los valores que minimizan:

$$Q_n(\beta_1, \beta_2, \lambda) = \sum_{t=2}^n (X_t - E[X_t/(\epsilon_{t-1}, \epsilon_{t-2})])^2 \quad (2.53)$$

con respecto a $(\beta_1, \beta_2, \lambda)$.

Derivando la ecuación (2.53) con respecto a β_1 , β_2 y λ , se tiene:

$$\hat{\beta}_1 = \frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]}{\left[\frac{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] - \left[\frac{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]} \quad (2.54)$$

$$\widehat{\beta}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] - \left[\frac{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right]} \quad (2.55)$$

y

$$\widehat{\lambda} = \bar{X}_t - \widehat{\beta}_1 \bar{\epsilon}_{t-1} - \widehat{\beta}_2 \bar{\epsilon}_{t-2} \quad (2.56)$$

Sustituyendo $\widehat{\beta}_1$ y $\widehat{\beta}_2$, respectivamente.

En efecto:

De la ecuación (2.53) y desarrollando el cuadrado, se tiene:

$$\begin{aligned} Q_n(\beta_1, \beta_2, \lambda) &= \sum_{t=2}^n (X_t - E[X_t / (\epsilon_{t-1}, \epsilon_{t-2})])^2 \\ &= \sum_{t=2}^n (X_t - (\beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \lambda))^2 \\ &= \sum_{t=2}^n (X_t^2 - 2X_t(\beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \lambda) + (\beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \lambda)^2) \\ &= \sum_{t=2}^n (X_t^2 - 2\beta_1 X_t \epsilon_{t-1} - 2\beta_2 X_t \epsilon_{t-2} - 2\lambda X_t + \beta_1^2 \epsilon_{t-1}^2 \\ &\quad + 2\beta_1 \beta_2 \epsilon_{t-1} \epsilon_{t-2} + \beta_2^2 \epsilon_{t-2}^2 + 2\beta_1 \lambda \epsilon_{t-1} + 2\beta_2 \lambda \epsilon_{t-2} + \lambda^2) \end{aligned}$$

derivando $Q_n(\beta_1, \beta_2, \lambda)$ con respecto a β_1 , β_2 y λ , tenemos:

$$\frac{\partial Q_n(\beta_1, \beta_2, \lambda)}{\partial \beta_1} = \sum_{t=2}^n (-2X_t \epsilon_{t-1} + 2\beta_1 \epsilon_{t-1}^2 + 2\beta_2 \epsilon_{t-1} \epsilon_{t-2} + 2\lambda \epsilon_{t-1}) = 0 \quad (2.57)$$

$$\frac{\partial Q_n(\beta_1, \beta_2, \lambda)}{\partial \beta_2} = \sum_{t=2}^n (-2X_t \epsilon_{t-2} + 2\beta_1 \epsilon_{t-1} \epsilon_{t-2} + 2\beta_2 \epsilon_{t-2}^2 + 2\lambda \epsilon_{t-2}) = 0 \quad (2.58)$$

$$\frac{\partial Q_n(\beta_1, \beta_2, \lambda)}{\partial \lambda} = \sum_{t=2}^n (-2X_t + 2\beta_1 \epsilon_{t-1} + 2\beta_2 \epsilon_{t-2} + 2\lambda) = 0 \quad (2.59)$$

En efecto:

La derivada de $Q_n(\beta_1, \beta_2, \lambda)$ con respecto a λ se tiene:

$$\begin{aligned}\widehat{\lambda} &= \frac{1}{n-1} \left[\sum_{t=2}^n X_t - \widehat{\beta}_1 \sum_{t=2}^n \epsilon_{t-1} - \widehat{\beta}_2 \sum_{t=2}^n \epsilon_{t-2} \right] \\ \widehat{\lambda} &= \overline{X}_t - \widehat{\beta}_1 \overline{\epsilon}_{t-1} - \widehat{\beta}_2 \overline{\epsilon}_{t-2} \quad \blacksquare\end{aligned}$$

sustituyendo el estimador $\widehat{\lambda}$ en la ecuación (2.57) tenemos para $\widehat{\beta}_1$:

$$\begin{aligned}\frac{\partial Q_n(\beta_1, \beta_2, \lambda)}{\partial \beta_1} &= \sum_{t=2}^n (-2X_t \epsilon_{t-1} + 2\beta_1 \epsilon_{t-1}^2 + 2\beta_2 \epsilon_{t-1} \epsilon_{t-2} + 2\lambda \epsilon_{t-1}) \\ &= -\sum_{t=2}^n X_t \epsilon_{t-1} + \widehat{\beta}_1 \sum_{t=2}^n \epsilon_{t-1}^2 + \widehat{\beta}_2 \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} + \widehat{\lambda} \sum_{t=2}^n \epsilon_{t-1} \\ &= -\sum_{t=2}^n X_t \epsilon_{t-1} + \widehat{\beta}_1 \sum_{t=2}^n \epsilon_{t-1}^2 + \widehat{\beta}_2 \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \\ &\quad + \left(\overline{X}_t - \widehat{\beta}_1 \overline{\epsilon}_{t-1} - \widehat{\beta}_2 \overline{\epsilon}_{t-2} \right) \sum_{t=2}^n \epsilon_{t-1} = 0\end{aligned}$$

despejando $\widehat{\beta}_1$ de este último resultado:

$$\begin{aligned}\widehat{\beta}_1 &= \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-1} - \overline{X}_t \sum_{t=2}^n \epsilon_{t-1} \right] + \widehat{\beta}_2 \left[\overline{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \overline{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]} \\ &= \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-1} - \overline{X}_t \sum_{t=2}^n \epsilon_{t-1} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \overline{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]} + \widehat{\beta}_2 \frac{\left[\overline{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \overline{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]}\end{aligned}\tag{2.60}$$

sustituyendo el estimador de λ en la ecuación (2.58) tenemos para $\widehat{\beta}_2$:

$$\begin{aligned}\frac{\partial Q_n(\beta_1, \beta_2, \lambda)}{\partial \beta_2} &= \sum_{t=2}^n (-2X_t \epsilon_{t-2} + 2\beta_1 \epsilon_{t-1} \epsilon_{t-2} + 2\beta_2 \epsilon_{t-2}^2 + 2\lambda \epsilon_{t-1}) \\ &= -\sum_{t=2}^n X_t \epsilon_{t-2} + \widehat{\beta}_1 \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} + \widehat{\beta}_2 \sum_{t=2}^n \epsilon_{t-2}^2 + \widehat{\lambda} \sum_{t=2}^n \epsilon_{t-1} \\ &= -\sum_{t=2}^n X_t \epsilon_{t-1} + \widehat{\beta}_1 \sum_{t=2}^n \epsilon_{t-1}^2 + \widehat{\beta}_2 \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \\ &\quad + \left(\overline{X}_t - \widehat{\beta}_1 \overline{\epsilon}_{t-1} - \widehat{\beta}_2 \overline{\epsilon}_{t-2} \right) \sum_{t=2}^n \epsilon_{t-1} = 0\end{aligned}$$

despejando $\hat{\beta}_2$ de este último resultado:

$$\begin{aligned}\hat{\beta}_2 &= \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2} \right] + \hat{\beta}_1 \left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} \\ &= \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} + \hat{\beta}_1 \frac{\left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]}\end{aligned}\quad (2.61)$$

hasta el momento, hemos logrado que las expresiones de $\hat{\beta}_1$ y $\hat{\beta}_2$, no dependan de $\hat{\lambda}$, pero dependen de $\hat{\beta}_2$ y $\hat{\beta}_1$ respectivamente, por lo que sustituyendo la ecuación (2.61) en la ecuación (2.60), tenemos:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]} + \frac{\left[\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]} \\ &\times \left(\frac{\left[\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} + \hat{\beta}_1 \frac{\left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} \right)\end{aligned}$$

Sabemos que:

$$\begin{aligned}\left[\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right] &= \left[\frac{1}{n-1} \sum_{t=2}^n \epsilon_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right] \\ &= \left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]\end{aligned}\quad (2.62)$$

Entonces, tenemos:

$$\hat{\beta}_1 \left(\frac{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} - \frac{\left[\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]^2}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]} \right) =$$

$$\left[\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1} \right] + \left[\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right] \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2} \right]}{\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right]}$$

Por lo que la expresión de $\hat{\beta}_1$ es:

$$\hat{\beta}_1 = \frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]}{\left[\frac{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] - \left[\frac{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]} \quad \blacksquare$$

Ahora, para encontrar el estimador de β_2 , se sigue un proceso similar al utilizado para β_1 , por lo que sustituyendo la ecuación (2.60) en la ecuación (2.61), tenemos:

$$\hat{\beta}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right] + \left[\frac{\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]}{\left(\frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right] + \hat{\beta}_2 \left[\frac{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right] + \left[\frac{\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}} \right]} \right)}$$

usando el resultado de la ecuación (2.62), tenemos:

$$\hat{\beta}_2 \left(\left[\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2} \right] - \frac{\left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right]^2}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]} \right) = \left[\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2} \right] + \left[\bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-2} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2} \right] \frac{\left[\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1} \right]}{\left[\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1} \right]}$$

Por lo que la expresión de $\hat{\beta}_2$ es:

$$\hat{\beta}_2 = \frac{\left[\frac{\sum_{t=2}^n X_t \epsilon_{t-2} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-2}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] + \left[\frac{\sum_{t=2}^n X_t \epsilon_{t-1} - \bar{X}_t \sum_{t=2}^n \epsilon_{t-1}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right]}{\left[\frac{\sum_{t=2}^n \epsilon_{t-2}^2 - \bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-2}}{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}} \right] - \left[\frac{\bar{\epsilon}_{t-2} \sum_{t=2}^n \epsilon_{t-1} - \sum_{t=2}^n \epsilon_{t-1} \epsilon_{t-2}}{\sum_{t=2}^n \epsilon_{t-1}^2 - \bar{\epsilon}_{t-1} \sum_{t=2}^n \epsilon_{t-1}} \right]} \quad \blacksquare$$

Finalmente, sustituyendo las expresiones encontradas para $\hat{\beta}_1$ y $\hat{\beta}_2$ en la ecuación (2.59), se obtiene el estimador de $\hat{\lambda}$. \blacksquare

Se observa que las expresiones encontradas para $\hat{\beta}_1$ y $\hat{\beta}_2$ son similarmente complejas, a las expresiones encontradas para $\hat{\alpha}_1$ y $\hat{\alpha}_2$ para un modelo INAR(2), entonces, al igual que para un modelo INAR(2) la complejidad de los estimadores de un modelo INMA(2), aumenta con el orden del modelo.

2.6. Estimación de Parámetros del Modelo INMA(q) con marginal de Poisson.

Recordemos que el modelo INMA(q) está dado por:

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

2.6.1. Estimadores Mínimos Cuadrados Condicionales.

La generalización del proceso de estimación de parámetros del modelo INMA(q), es similar a la estimación de un modelo INAR(p) estudiado en el apartado anterior, y representado por Du y Li (1991) y Funes (2001).

En el modelo INMA(q) se tiene que dado $(\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p})$, X_t sigue siendo variable aleatoria. Por lo que:

$$\begin{aligned} E(X_t / (\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p})) &= \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \dots + \beta_p \epsilon_{t-p} + \lambda \\ &= g(\theta, (\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p})) \end{aligned}$$

donde $\theta = (\beta_1, \beta_2, \dots, \beta_p, \lambda)^T$ es el conjunto de parámetros a estimar. 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:

$$\sum_{t=p}^n (X_t - E(X_t / (\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-p})))^2 \quad (2.63)$$

Diferenciando la ecuación (2.63) con respecto a θ_i con $i = 1, 2, \dots, (p+1)$ e igualando a cero se obtienen los estimadores de mínimos cuadrados condicionales de $\hat{\theta}_i$. ■

Capítulo 3

Estudio de simulación y aplicaciones reales.

3.1. Simulación del modelo INAR

Los resultados de la simulación para el modelo INAR(1) con $\lambda = 1$ se presentan en las Tablas 3.1, 3.2 y 3.3 las cuales muestran el método de Yule-Walker (YW), Mínimos Cuadrados Condicionales (MMC) y Máxima Verosimilitud Condicional (MV), respectivamente.

En la Tabla 3.1 (método de Yule-Walker) se observa que el sesgo $\hat{\alpha}$ es positivo cuando $\alpha = 0.1$ y $n = 150$, el resto de los valores simulados presentan sesgo negativo (96 %), es decir, se ha obtenido una subestimación del parámetro α ; mientras que el sesgo de $\hat{\lambda}$ presenta un comportamiento opuesto al sesgo de $\hat{\alpha}$, ya que presenta dos valores negativos y precisamente uno de los valores negativos es cuando $\alpha = 0.1$ y $n = 150$, el segundo valor negativo es cuando $\alpha = 0.1$ y $n = 200$, para el resto de los valores de α y n el sesgo de $\hat{\lambda}$ es positivo (92 %), en este caso se ha obtenido una sobreestimación del parámetro λ . Por otra parte, se observa que el error cuadrático medio de $\hat{\lambda}$ es siempre mayor que el de $\hat{\alpha}$.

Cuadro 3.1: INAR(1). Yule-Walker

Sesgo y Error Cuadrático Medio de $\hat{\alpha}$ y $\hat{\lambda}$ ($\lambda = 1$)					
n	α	sesgo de $\hat{\alpha}$	sesgo de $\hat{\lambda}$	ECM($\hat{\alpha}$)	ECM($\hat{\lambda}$)
		YW	YW	YW	YW
50	0.1	-0.03281	0.03298	0.02054	0.04133
	0.3	-0.04965	0.08187	0.01876	0.05482
	0.5	-0.06223	0.14951	0.02429	0.12542
	0.7	-0.07362	0.19516	0.01649	0.19228
	0.9	-0.06384	0.46321	0.00870	0.53079
75	0.1	-0.02136	0.03135	0.01064	0.02537
	0.3	-0.04385	0.05100	0.01532	0.04863
	0.5	-0.04025	0.08503	0.01703	0.06948
	0.7	-0.04248	0.13965	0.01059	0.11030
	0.9	-0.05580	0.47448	0.00543	0.39320
100	0.1	-0.02295	0.03505	0.01253	0.02779
	0.3	-0.02942	0.02315	0.00924	0.02235
	0.5	-0.03262	0.05487	0.01097	0.04727
	0.7	-0.04042	0.14215	0.00756	0.10778
	0.9	-0.04465	0.42451	0.00479	0.44116
150	0.1	0.00638	-0.01194	0.00731	0.01641
	0.3	-0.02467	0.03086	0.00713	0.01750
	0.5	-0.02676	0.03946	0.00622	0.02704
	0.7	-0.02914	0.09978	0.00473	0.05124
	0.9	-0.03325	0.30900	0.00269	0.23805
200	0.1	-0.00740	-0.00868	0.00486	0.00947
	0.3	-0.00555	0.00777	0.00586	0.01450
	0.5	-0.01750	0.03402	0.00528	0.02516
	0.7	-0.02593	0.08755	0.00323	0.03844
	0.9	-0.02657	0.25127	0.00203	0.18127

Ahora, de la Tabla 3.2 (método de Mínimos Cuadrados Condicionales) se observa un comportamiento similar para $\hat{\alpha}$ y $\hat{\lambda}$ en cuanto al sesgo y error cuadrático medio, comparado con método de Yule-Walker.

Cuadro 3.2: INAR(1). Mínimos Cuadrados Condicionales

Sesgo y Error Cuadrático Medio de $\hat{\alpha}$ y $\hat{\lambda}$ ($\lambda = 1$)					
n	α	sesgo de $\hat{\alpha}$	sesgo de $\hat{\lambda}$	ECM($\hat{\alpha}$)	ECM($\hat{\lambda}$)
		MCC	MCC	MCC	MCC
50	0.1	-0.02803	0.02789	0.02067	0.04162
	0.3	-0.04148	0.07047	0.01800	0.05257
	0.5	-0.05240	0.13060	0.02290	0.11830
	0.7	-0.05159	0.12870	0.01361	0.16096
	0.9	-0.03701	0.25920	0.00520	0.29875
75	0.1	-0.01948	0.02924	0.01080	0.02542
	0.3	-0.03978	0.04536	0.01530	0.04836
	0.5	-0.03343	0.07178	0.01627	0.06606
	0.7	-0.02900	0.09728	0.00966	0.09977
	0.9	-0.03679	0.31630	0.00342	0.25045
100	0.1	-0.02219	0.03429	0.01271	0.02817
	0.3	-0.02541	0.01740	0.00917	0.02169
	0.5	-0.02950	0.04876	0.01072	0.04616
	0.7	-0.03355	0.11852	0.00669	0.09216
	0.9	-0.03127	0.30024	0.00341	0.31134
150	0.1	0.00661	-0.01216	0.00743	0.01656
	0.3	-0.02219	0.02748	0.00701	0.01763
	0.5	-0.02348	0.03302	0.00621	0.02695
	0.7	-0.02437	0.08097	0.00392	0.04949
	0.9	-0.02502	0.23192	0.00200	0.17458
200	0.1	-0.00717	-0.00895	0.00475	0.00932
	0.3	-0.00364	0.00498	0.00587	0.01432
	0.5	-0.01540	0.02985	0.00512	0.02434
	0.7	-0.02142	0.07263	0.00297	0.03517
	0.9	-0.02081	0.19784	0.00160	0.14588

Por último, en la Tabla 3.3 (método de Máxima Verosimilitud) se observa que el sesgo de $\hat{\alpha}$ es negativo en todos los casos (100%), es decir, se ha obtenido una subestimación del parámetro α ; mientras que el sesgo de $\hat{\lambda}$ presenta un comportamiento opuesto al sesgo de $\hat{\alpha}$, ya que presenta un solo valor negativo, precisamente este valor es cuando $\alpha = 0.1$ y $n = 200$, para el resto de los valores de α y n el sesgo de $\hat{\lambda}$ es positivo (96%), en este caso se ha obtenido una sobreestimación del parámetro λ . Por otra parte, el error cuadrático medio de $\hat{\lambda}$ es siempre mayor que el de $\hat{\alpha}$ únicamente para $\alpha = 0.9$, las demás combinaciones, tienen comportamiento similar para los estimadores.

Cuadro 3.3: INAR(1). Máxima Verosimilitud

Sesgo y Error Cuadrático Medio de $\hat{\alpha}$ y $\hat{\lambda}$ ($\lambda = 1$)					
n	α	sesgo de $\hat{\alpha}$	sesgo de $\hat{\lambda}$	ECM($\hat{\alpha}$)	ECM($\hat{\lambda}$)
		MV	MV	MV	MV
50	0.1	-0.01606	0.04313	0.00651	0.04233
	0.3	-0.02046	0.03100	0.00680	0.00818
	0.5	-0.02415	0.04530	0.01384	0.00418
	0.7	-0.02698	0.01729	0.00664	0.00587
	0.9	-0.02262	0.12186	0.00171	0.05793
75	0.1	-0.01403	0.01631	0.00623	0.00461
	0.3	-0.01898	0.00397	0.00693	0.00639
	0.5	-0.02298	0.04127	0.00577	0.00507
	0.7	-0.02571	0.07512	0.00440	0.02555
	0.9	-0.02681	0.23716	0.00194	0.20609
100	0.1	-0.01408	0.01879	0.00348	0.00269
	0.3	-0.01860	0.00826	0.00232	0.00848
	0.5	-0.02192	0.03063	0.00271	0.00836
	0.7	-0.02405	0.06585	0.00507	0.01933
	0.9	-0.00252	0.01232	0.00053	0.00762
150	0.1	-0.01272	0.00612	0.00155	0.00305
	0.3	-0.01762	0.01990	0.00114	0.00376
	0.5	-0.02042	0.02517	0.00150	0.00481
	0.7	-0.02394	0.06744	0.00342	0.01638
	0.9	-0.02401	0.19656	0.00221	0.11680
200	0.1	-0.01109	-0.00733	0.00461	0.00355
	0.3	-0.02251	0.03086	0.00317	0.00645
	0.5	-0.01904	0.03508	0.00140	0.00462
	0.7	-0.02110	0.06839	0.00209	0.02108
	0.9	-0.02386	0.23137	0.00213	0.22335

Con el fin de comparar la eficiencia de los tres métodos de estimación respecto a los resultados obtenidos, podemos ordenarlos de la siguiente forma, el primero puede ser el método de Máxima Verosimilitud Condicional, seguido del método de Mínimos Cuadrados Condicionales y finalmente el método de Yule-Walker.

Los resultados de la simulación para el modelo INAR(2), se presentan en las Tablas 3.4 y 3.5, las cuales muestran el método de Yule-Walker (YW) y Mínimos Cuadrados Condicionales (MMC). El experimento se realiza para modelos de tamaño, $n = 50, 75, 100, 150, 200$ y diferentes valores de $\alpha_1 = 0.1, 0.3, 0.5, 0.7, 0.8$ y ($\lambda = 1$ y $\alpha_2 = 0.1$).

Cuadro 3.4: INAR(2). Yule-Walker

Sesgo y Error Cuadrático Medio de $\widehat{\alpha}_1$, $\widehat{\alpha}_2$ y $\widehat{\lambda}$ ($\lambda = 1$, $\alpha_2 = 0,1$)							
n	α_1	sesgo de $\widehat{\alpha}_1$	sesgo de $\widehat{\alpha}_2$	sesgo de $\widehat{\lambda}$	ECM($\widehat{\alpha}_1$)	ECM($\widehat{\alpha}_2$)	ECM($\widehat{\lambda}$)
		YW	YW	YW	YW	YW	YW
50	0.1	-0.03227	-0.05139	0.08807	0.02739	0.02146	0.09359
	0.3	-0.04395	-0.05699	0.11920	0.02395	0.02682	0.11916
	0.5	-0.02385	-0.06508	0.19527	0.02199	0.02384	0.15760
	0.7	-0.01007	-0.07811	0.34554	0.02804	0.03275	0.42515
	0.8	0.01491	-0.08915	0.51916	0.02248	0.02779	0.52552
75	0.1	-0.01207	-0.04072	0.05690	0.01386	0.01556	0.05610
	0.3	-0.01336	-0.04337	0.09110	0.01536	0.01440	0.07901
	0.5	-0.02048	-0.05212	0.16780	0.01704	0.01554	0.11499
	0.7	-0.04170	0.00287	0.16386	0.01697	0.01546	0.19660
	0.8	-0.01198	-0.03453	0.37996	0.01493	0.01360	0.41073
100	0.1	-0.02384	-0.03554	0.06015	0.01150	0.01280	0.03940
	0.3	-0.01913	-0.04290	0.05610	0.01588	0.01762	0.05295
	0.5	-0.00725	-0.02238	0.06573	0.01087	0.01110	0.06600
	0.7	-0.02109	-0.02238	0.17489	0.01140	0.01000	0.13305
	0.8	0.00107	-0.03632	0.30443	0.00888	0.01066	0.27302
150	0.1	0.00510	-0.01823	0.01960	0.00937	0.00717	0.02747
	0.3	-0.00362	-0.02114	0.04294	0.00868	0.00789	0.03500
	0.5	-0.00608	-0.01252	0.03978	0.00820	0.00814	0.04061
	0.7	-0.00450	-0.02972	0.16279	0.00705	0.00710	0.15977
	0.8	0.01153	-0.03461	0.19379	0.00617	0.00719	0.17976
200	0.1	-0.00021	-0.01883	0.02244	0.00660	0.00495	0.01901
	0.3	0.00201	-0.00774	0.00906	0.00561	0.00523	0.02067
	0.5	-0.00974	-0.01481	0.05040	0.00554	0.00590	0.03582
	0.7	-0.01945	-0.00352	0.10742	0.00559	0.00443	0.05941
	0.8	-0.00099	-0.02015	0.19045	0.00462	0.00504	0.12763

En la Tabla 3.4 (método de Yule-Walker) se observa que el sesgo de $\widehat{\alpha}_1$ es negativo en un (80%) de los casos, es decir, se ha obtenido una subestimación del parámetro α_1 ; mientras que el sesgo de $\widehat{\alpha}_2$ es positivo únicamente para $n = 75$ y $\alpha_1 = 0.7$, el resto que representa un (96%) de negatividad muestra el que parámetro α_2 ha sido subestimado; finalmente el sesgo de $\widehat{\lambda}$ es siempre positivo, presentando una sobreestimación del parámetro λ . Respecto al error cuadrático medio de cada parámetro podemos decir que este disminuye a medida que aumenta n para α_1 .

Cuadro 3.5: INAR(2). Mínimos Cuadrados Condicionales

Sesgo y Error Cuadrático Medio de $\widehat{\alpha}_1$, $\widehat{\alpha}_2$ y $\widehat{\lambda}$ ($\lambda = 1$, $\alpha_2 = 0,1$)							
n	α_1	sesgo de $\widehat{\alpha}_1$	sesgo de $\widehat{\alpha}_2$	sesgo de $\widehat{\lambda}$	ECM($\widehat{\alpha}_1$)	ECM($\widehat{\alpha}_2$)	ECM($\widehat{\lambda}$)
		MCC	MCC	MCC	MCC	MCC	MCC
50	0.1	-0.03296	-0.04170	0.07712	0.02704	0.02076	0.08860
	0.3	-0.04415	-0.04963	0.10882	0.02451	0.02536	0.11670
	0.5	-0.02813	-0.05358	0.17807	0.02196	0.02129	0.13892
	0.7	-0.03315	-0.04460	0.30722	0.03059	0.02746	0.36170
	0.8	-0.02744	-0.03271	0.43048	0.02739	0.02325	0.42351
75	0.1	-0.01142	-0.03995	0.05495	0.01389	0.01440	0.05409
	0.3	-0.01441	-0.03982	0.08708	0.01506	0.01392	0.07795
	0.5	-0.02315	-0.04510	0.15765	0.01657	0.01485	0.11101
	0.7	-0.04923	0.01347	0.15398	0.01759	0.01481	0.17658
	0.8	-0.04944	0.00711	0.35633	0.02036	0.01505	0.35733
100	0.1	-0.02370	-0.03440	0.05832	0.01136	0.01257	0.03862
	0.3	-0.01968	-0.04332	0.05776	0.01600	0.01723	0.05281
	0.5	-0.00836	-0.02016	0.06443	0.01058	0.01030	0.06466
	0.7	-0.03111	-0.01080	0.16959	0.01251	0.01054	0.12434
	0.8	-0.01971	-0.01470	0.30411	0.01157	0.01117	0.25620
150	0.1	0.00476	-0.01688	0.01835	0.00933	0.00700	0.02720
	0.3	-0.00362	-0.01931	0.04026	0.00878	0.00807	0.03638
	0.5	-0.00637	-0.01084	0.03695	0.00818	0.00808	0.03957
	0.7	-0.00926	-0.02550	0.16577	0.00741	0.00712	0.15781
	0.8	-0.00172	-0.02169	0.19835	0.00703	0.00723	0.17683
200	0.1	-0.00032	-0.01822	0.02181	0.00657	0.00488	0.01879
	0.3	0.00168	-0.00765	0.00958	0.00563	0.00521	0.02073
	0.5	-0.01015	-0.01528	0.05272	0.00560	0.00595	0.03615
	0.7	-0.02317	-0.00019	0.10962	0.00598	0.00447	0.05820
	0.8	-0.01167	-0.00989	0.19485	0.00497	0.00487	0.12497

En la Tabla 3.5 (método Mínimos Cuadrados Condicionales) se observa que el sesgo de $\widehat{\alpha}_1$ es negativo en un (92 %), es decir, se ha obtenido una subestimación del parámetro α_1 ; este mismo comportamiento muestra α_2 ; finalmente el sesgo de λ es siempre positivo, presentando una sobreestimación del parámetro. Respecto al error cuadrático medio de cada parámetro se observa un comportamiento similar al de la Tabla 3.4.

3.2. Simulación del modelo INMA

Los resultados de la simulación para el modelo INMA(1) con $\lambda = 1$, se presentan en la Tabla 3.6 la cual muestra el método de Mínimos Cuadrados Condicionales (MMC).

Cuadro 3.6: INMA(1). Mínimos Cuadrados Condicionales

Sesgo y Error Cuadrático Medio de $\hat{\beta}$ y $\hat{\lambda}$ ($\lambda = 1$)					
n	β	sesgo de $\hat{\beta}$	sesgo de $\hat{\lambda}$	ECM($\hat{\beta}$)	ECM($\hat{\lambda}$)
		MCC	MCC	MCC	MCC
50	0.1	-0.02328	0.01797	0.02324	0.03552
	0.3	-0.04041	0.03728	0.03056	0.04393
	0.5	-0.01359	0.00951	0.02971	0.03499
	0.7	0.00993	-0.00144	0.03111	0.03819
	0.9	-0.04264	0.01203	0.02529	0.03125
75	0.1	0.00387	-0.00723	0.01679	0.02065
	0.3	-0.02945	0.04884	0.02010	0.03027
	0.5	-0.01601	0.02518	0.01803	0.01976
	0.7	-0.01034	-0.01164	0.01965	0.03011
	0.9	-0.00704	0.00016	0.01434	0.02048
100	0.1	-0.01474	0.01375	0.01190	0.02476
	0.3	-0.03338	0.02676	0.01182	0.01980
	0.5	-0.00986	-0.00534	0.01730	0.02082
	0.7	-0.01260	0.00065	0.01443	0.01855
	0.9	0.00839	0.00468	0.01092	0.02273
150	0.1	0.00081	-0.00055	0.00837	0.01398
	0.3	-0.00875	0.02083	0.00985	0.01808
	0.5	-0.00775	0.00776	0.00972	0.01477
	0.7	-0.00992	-0.00727	0.00770	0.01116
	0.9	-0.00354	0.01894	0.00601	0.01096
200	0.1	-0.00168	-0.00379	0.00717	0.00930
	0.3	0.00510	-0.01064	0.00883	0.01224
	0.5	0.00915	-0.00555	0.00599	0.01119
	0.7	-0.00260	-0.00585	0.00717	0.01059
	0.9	-0.00505	0.00180	0.00524	0.01038

En la Tabla 3.6 (método Mínimo Cuadrado Condicionales) se observa que el sesgo de $\hat{\beta}$ es positivo en seis de los casos, el resto de los valores simulados presentan sesgo negativo (76%), es decir, se ha obtenido una subestimación del parámetro β ; mientras que el sesgo de $\hat{\lambda}$ presenta un comportamiento no muy definido, ya que diez de los valores simulados son negativos (40%), en este caso se ha obtenido una leve sobreestimación del parámetro λ . Por

otra parte, se observa que el error cuadrático medio de $\hat{\lambda}$ es siempre mayor que el de $\hat{\beta}$.

Los resultados de la simulación para el modelo INMA(2) se presentan en las Tabla 3.7 la cual muestra el método de Mínimos Cuadrados Condicionales (MMC). El experimento se realiza para modelos de tamaño, $n = 50, 75, 100, 150, 200$ y diferentes valores de $\beta_1 = 0.1, 0.3, 0.5, 0.7, 0.8$ y ($\lambda = 1$ y $\beta_2 = 0.1$).

Cuadro 3.7: INMA(2). Mínimos Cuadrados Condicionales

Sesgo y Error Cuadrático Medio de $\hat{\beta}_1, \hat{\beta}_2$ y $\hat{\lambda}$ ($\lambda = 1, \beta_2 = 0,1$)							
n	β_1	sesgo de $\hat{\beta}_1$	sesgo de $\hat{\beta}_2$	sesgo de $\hat{\lambda}$	ECM($\hat{\beta}_1$)	ECM($\hat{\beta}_2$)	ECM($\hat{\lambda}$)
		MCC	MCC	MCC	MCC	MCC	MCC
50	0.1	-0.01533	-0.05219	0.09322	0.02800	0.02903	0.08733
	0.3	0.00428	-0.02959	0.01525	0.03976	0.02684	0.06951
	0.5	-0.03327	-0.04527	0.06873	0.03208	0.02862	0.07899
	0.7	-0.02469	-0.03207	0.02326	0.02707	0.02975	0.08185
	0.8	-0.04749	-0.05613	0.13218	0.02520	0.03072	0.09976
75	0.1	-0.02067	-0.03185	0.05422	0.01588	0.01577	0.04282
	0.3	-0.02270	-0.03518	0.06714	0.02181	0.01862	0.06190
	0.5	-0.00808	-0.03224	0.00715	0.02269	0.01682	0.04431
	0.7	-0.01210	-0.02080	0.00944	0.02558	0.01695	0.03868
	0.8	-0.02233	-0.02161	0.00713	0.02703	0.01797	0.04604
100	0.1	-0.00721	-0.00813	0.03721	0.01205	0.01233	0.03516
	0.3	-0.00455	-0.01673	0.03408	0.01199	0.01515	0.03698
	0.5	-0.01207	-0.01899	0.00945	0.01236	0.01413	0.03315
	0.7	-0.01154	-0.03367	0.04936	0.01560	0.01893	0.04637
	0.8	-0.02592	-0.01444	0.02964	0.01370	0.01569	0.03448
150	0.1	-0.00580	-0.01399	0.01996	0.00649	0.00776	0.02029
	0.3	-0.00182	-0.02370	0.03776	0.01134	0.00841	0.02784
	0.5	-0.01613	-0.00487	0.01068	0.01246	0.01134	0.02842
	0.7	-0.01807	-0.01790	0.03452	0.00936	0.00700	0.02353
	0.8	0.00395	-0.01296	0.01205	0.00873	0.00970	0.02620
200	0.1	-0.00792	-0.01037	0.01073	0.00609	0.00752	0.01653
	0.3	-0.01775	-0.00190	0.03080	0.00921	0.00685	0.01965
	0.5	-0.00151	-0.00341	-0.00256	0.00705	0.00778	0.01425
	0.7	-0.02067	-0.00412	0.02793	0.00911	0.00573	0.01763
	0.8	0.00291	-0.00287	-0.00598	0.00666	0.00688	0.01815

En la Tabla 3.7 se observa que el sesgo de β_1 es positivo en tres de los casos, obteniendo una subestimación del (88%) de los casos, mientras que el sesgo de β_2 presenta un (100%) de subestimación; finalmente el sesgo de λ es negativo en dos de los casos, para $n = 200$ con

$\beta_1 = 0.5$ y 0.8 , obteniendo un (92%) de sobreestimación del parámetro. Respecto al error cuadrático medio no se identifica un comportamiento de cada parámetro, los dos estimadores (β_1 y β_2) tienen ECM muy similares.

3.3. Aplicación de datos reales.

3.3.1. Casos de Dengue.

En esta sección trataremos de ajustar un modelo a la serie: Número de casos diarios de Dengue durante el 2005. La Figura 3.1 presenta el comportamiento de dicha serie.

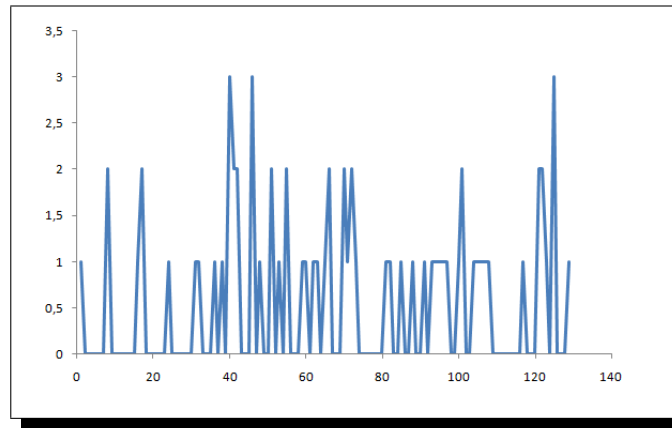


Figura 3.1: Número de casos diarios de Dengue durante el 2005. *Fuente:* Hospital I.S.S.S. de Amatepec, Soyapango.

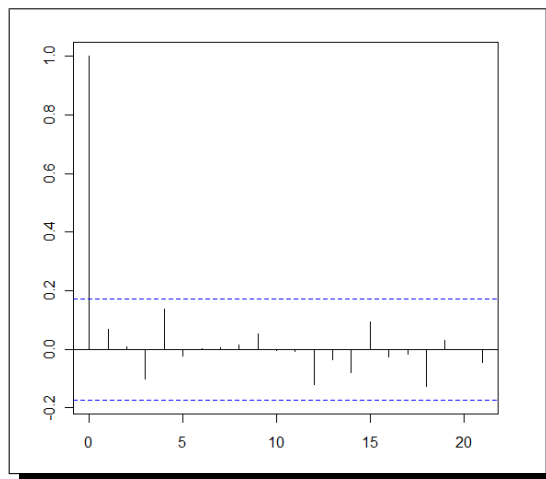


Figura 3.2: Función de autocorrelación para la serie temporal Dengue.

La función de autocorrelación de los casos de Dengue se muestra en la Figura 3.2, el posible modelo para esta serie se puede definir como un ruido aleatorio de la siguiente forma:

$$x_t = \mu + \epsilon_t \quad \text{con} \quad \mu = 0.5271 \quad \text{la media de la serie.} \quad (3.1)$$

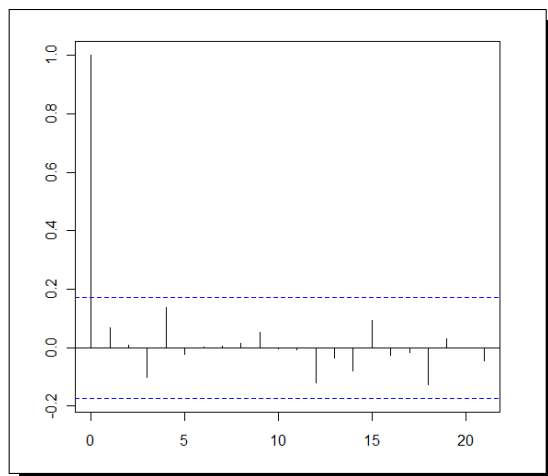


Figura 3.3: Función de autocorrelación para los errores del modelo (3.1).

La Figura 3.3 muestra la función de autocorrelación de los errores para el modelo ajustado a la serie Dengue, esto muestra un comportamiento de variables aleatorias i.i.d. por lo tanto, la serie esta dada por la ecuación (3.1), donde $\epsilon_t \sim \mathbb{P}_0(0.004)$.

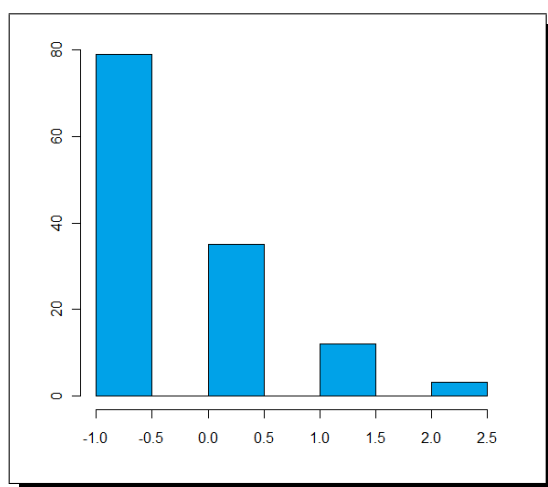


Figura 3.4: Histograma de los errores de la ecuación (3.1).

La Figura 3.4 muestra el histograma de los errores de la serie Denge ajustada a la ecuación (3.1), ésta figura muestra que los errores presentan un comportamiento de un distribución de Poisson. El modelo ajustado es el siguiente:

$$x_t = 0.5271 + \epsilon_t$$

donde ϵ_t tiene una distribución de Poisson de media 0.004.

3.3.2. Número de Partos.

En esta sección trataremos de ajustar un modelo a la serie: Número de partos diarios durante el 2006. La Figura 3.5 presenta el comportamiento de dicha serie.

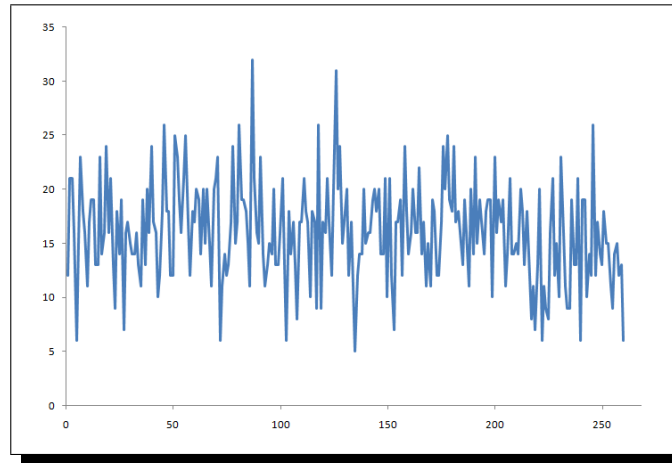


Figura 3.5: Número de partos diarios durante el 2006. *Fuente:* Hospital I.S.S.S. Amatepec, Soyapango.

La función de autocorrelación de la serie Partos se muestra en la Figura 3.6, se observa que dos correlaciones superan los límites, por lo que podría suponerse un comportamiento de segundo orden, pero entre ellas hay un decrecimiento rápido hacia cero. Entonces, de esta figura notamos que la serie puede ser modelada con un proceso INAR(5), definido de la siguiente forma:

$$x_t = \alpha_5 * x_{t-5} + \epsilon_t \quad (3.2)$$

Realizaremos las estimaciones de los parámetros del modelo (3.2) asumiendo que $\epsilon_t \sim \mathbb{P}_0(\lambda)$ y utilizando los métodos de Yule-Walker (YW), Mínimos Cuadrados Condicionales (MCC) y Máxima Verosimilitud Condicional (MV), los resultados se muestran en la tabla 3.8.

Cuadro 3.8: Estimación de parámetros del modelo INAR(5), ajustado al modelo (3.2)

Método	$\hat{\alpha}$	$\hat{\lambda}$
YW	0.18181	13.15817
MCC	0.20714	12.75790
MV	0.33782	10.63174

Entonces, las ecuaciones de la serie Partos, según la Tabla 3.8 para los métodos Mínimos

Cuadrados Condicionales y Máxima Verosimilitud son respectivamente:

$$x_t = 0.20714 * x_{t-5} + \epsilon_t \text{ con } \epsilon_t \sim \mathbb{P}_0(12.75790) \quad (3.3)$$

$$x_t = 0.33782 * x_{t-5} + \epsilon_t \text{ con } \epsilon_t \sim \mathbb{P}_0(10.63174) \quad (3.4)$$

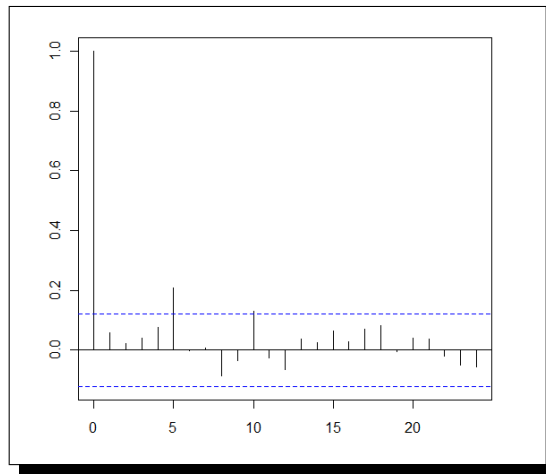


Figura 3.6: Función de autocorrelación para la serie temporal Partos.

Con la idea de conocer como se comportan los errores de la serie Partos según las ecuaciones 3.3 y 3.4, se presenta el histograma y la función de autocorrelación, respectivamente.

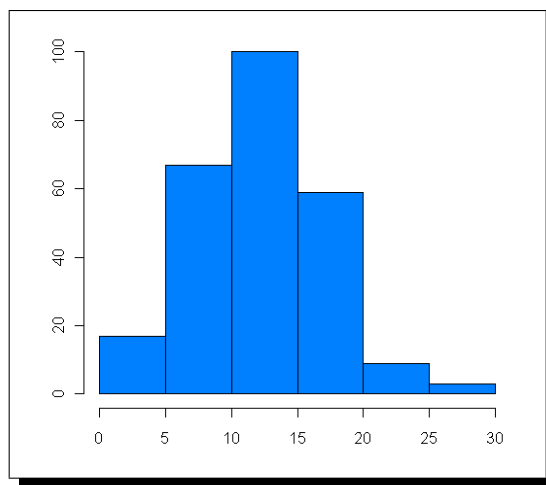


Figura 3.7: Histograma de los errores de la ecuación (3.3).

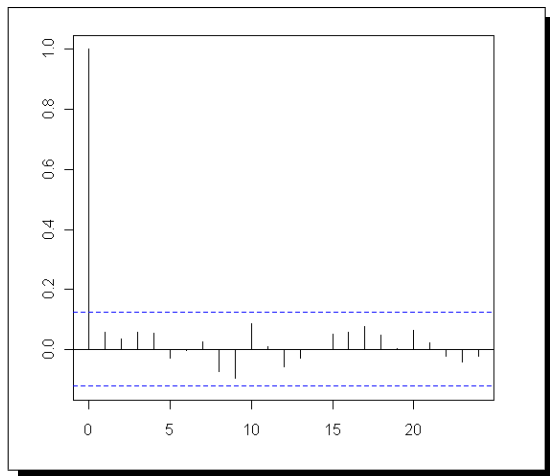


Figura 3.8: Función de autocorrelación de los errores de la ecuación (3.3).

La figura 3.7, muestra que los errores del modelo ajustado a la serie Partos aplicando el método de Mínimos Cuadrados Condicionales presenta un comportamiento de una distribución de Poisson, además se observa en la figura 3.8, que estos errores presentan un comportamiento de variable aleatoria i.i.d. entonces la serie está dada por la ecuación (3.3), donde $\epsilon_t \sim \mathbb{P}_0(12.76)$. El modelo ajustado al número de nacimientos en el hospital utilizando el método de Mínimos Cuadrados Condicionales esta dado por:

$$x_t = 0.20714 * x_{t-5} + \epsilon_t$$

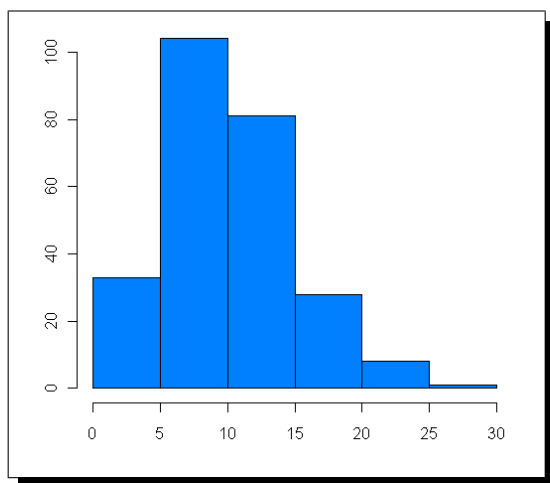


Figura 3.9: Histograma de los errores de la ecuación (3.4).

Finalmente, la figura 3.9, muestra que los errores del modelo ajustado a la serie Partos aplicando el método de Máxima Verosimilitud también presenta un comportamiento de una distribución de Poisson, además la figura 3.10, se observa que los errores presentan un comportamiento de variable aleatoria i.i.d. entonces la serie está dada por la ecuación (3.4), donde $\epsilon_t \sim \mathbb{P}_0(10.56)$. El modelo ajustado al número de nacimientos en el hospital utilizando el método de Máxima Verosimilitud esta dado por:

$$x_t = 0.33782 * x_{t-5} + \epsilon_t$$

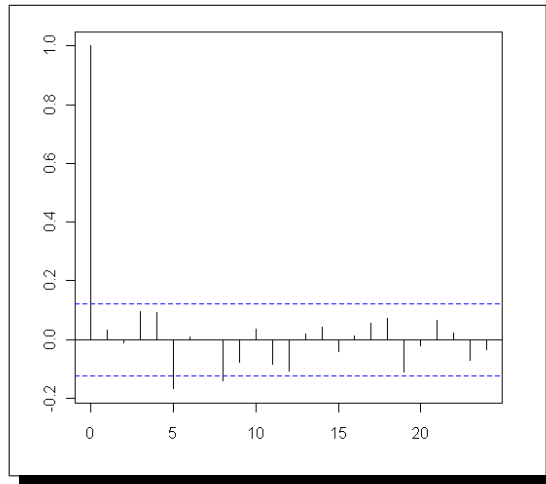


Figura 3.10: Función de Autocorrelación de los errores de la ecuación (3.4).

Sabemos que la estimación de λ usando el método de Mínimos Cuadrados Condicionales y el método de Máxima Verosimilitud son 12.75790 y 10.63174 respectivamente (ver tabla 3.8), éstos valores son muy próximos a la media del proceso de los errores, 12.76 y 10.56. Esto muestra que la simulación realizada usando los métodos mencionados anteriormente arroja buenos resultados.

Conclusiones

La función de autocorrelación de los modelos INARMA, es idéntica a la de los modelos ARMA estándar, a pesar que los primeros tienen media y varianza de una distribución discreta.

Al igual que para los modelos ARMA estándar, los procesos INARMA deben cumplir las condiciones de estacionariedad e invertibilidad.

Se verifica que para procesos estacionarios, la estimación de parámetros para un proceso INAR(1) usando el método de Yule-Walker y Mínimos Cuadrados Condicionales, se obtienen resultados similares.

Comparando los métodos de estimación en relación con los resultados de la simulación del proceso INAR(1), se observa que con el método de Máxima Verosimilitud Condicional se obtienen mejores resultados.

De la aplicación real del número de partos, se observa que se ajusta un modelo INAR(5), donde influye para la observación presente lo ocurrido cinco días antes más un proceso de innovación con distribución de Poisson.

Como futura línea de investigación se plantea:

Abordar el tema de definir la estructura y propiedades de un modelo mixto INARMA(p, q) para p y q mayores que la unidad.

Obtener los estimadores de máxima verosimilitud de un modelo INAR(p) para p mayor que la unidad.

Bibliografía

Al-Osh, M. A. and Alzaid, A. A. (1987) First-order integer valued autoregressive INAR(1) process: *Journal Time Series Analysis*, 8, 261-275.

Al-Osh, M. A. and Alzaid, A. A. (1990) An integer-valued pth autoregressive structure (INAR(p)): *Journal Appl. Probab.* 27, 314-324.

Aznar, Anatonio y Trívez, F. Javier (1993) *Métodos de predicción en Economía II*. 1º Edición. Editorial Ariel S.A. Barcelona.

Coleman, Rodney (1986) *PROCESOS ESTOCÁSTICOS*. Editorial Limusa.

David A. Sprott (1983) Estimating the parameters of a convolution by Maximun Likelihood: *Journal of the American Statistical Association*. 78, 457-460.

Funes Torres, Nerys (2001) *Modelos INARMA para Series Temporales de valores Enteros: Análisis, propiedades asintóticas y estimación*. Universidad Complutense de Madrid, España. (Tesis Doctoral).

G. Gauthier y Alain Latour (1994) Convergence forte de estimateurs des paramètres d'un processus GENAR(p): *Ann. Sci. Québec*. 18, 49-71.

J. Franke, T. Seligman (1993) Conditional maximun likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. 22, 311-330.

Klimko, Lawrence A. y Nelson, Paul I. (1978) ON CONDITIONAL LEAST SQUARES ESTIMATION FOR STOCHASTIC PROCESSES. *The Annals of Statistics*, 3, 629-642.

Marques da Silva, Isabel Maria (2005) Contributions to the analysis of discrete-valued time series. Universidade do Porto, Portugal. (Tesis Doctoral).

Park, YouSung y Kim, MyungJin (1997) Some Basic and Asymptotic Properties in INMA(q) Processes. 2, 155-170.

YouSung Park, Chan Wook Oh (1997) Some asymptotic properties in INAR(1) processes with Poisson marginals. *Statistical Papers*, 38, 287-302.